

187514

STIC-Biotech/ChemLib

From: Swope, Sheridan  
Sent: Thursday, April 27, 2006 4:33 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/825,632

For 10/825,632, pls search and interference search:

SID 2, oligo search, against the NT and AA databases.

SID 4, regular search, against the NT and AA databases.

SID 6, regular search, against the NT and AA databases.

SID 8, regular search, against the NT and AA databases.

*Not a tag of SID 2*  
Sheridan Swope, Ph.D.  
Primary Patent Examiner  
AU 1656/Recombinant Enzymes  
571-272-0943 (voice)  
E02B71 Remsen Bld (Office)  
E03C70 Remsen Bld (Mailbox)

CRFE

BEST AVAILABLE COPY

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_



**THIS PAGE BLANK (USPTO)**



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 187514**

**TO: Sheridan Swope**  
**Location: REM-2B71&3C70**  
**Art Unit: 1656**  
**Tuesday, May 09, 2006**  
**Case Serial Number: 10/825632**

**From: Kristine Hensle**  
**Location: Biotech-Chem Library**  
**REM-1B69**  
**Phone: (571) 272-4161**

**Kristine. Hensle@uspto.gov**

### **Search Notes**

Examiner Swope,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Kristine Hensle  
Librarian  
STIC Biotech/Chem Library  
(571)272-4161



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 04:50:59 ; Search time 15358 Seconds  
(without alignments)  
11547.830 Million cell updates/sec

Title: US-10-825-632-2  
Perfect score: 3120  
Sequence: 1 aagtgttaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: 

Searched: 5883141 seqs, 28421725653 residues

Word size: 1

Total number of hits satisfying chosen parameters: 11765110

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- GenEmbl:\*
- 1: gb\_ba:\*
  - 2: gb\_in:\*
  - 3: gb\_env:\*
  - 4: gb\_on:\*
  - 5: gb\_ov:\*
  - 6: gb\_pat:\*
  - 7: gb\_ph:\*
  - 8: gb\_pr:\*
  - 9: gb\_ro:\*
  - 10: gb\_sts:\*
  - 11: gb\_sy:\*
  - 12: gb\_un:\*
  - 13: gb\_vi:\*
  - 14: gb\_htg:\*
  - 15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3120	100.0	3120	6	AR651452 Sequence
2	3120	100.0	3127	8	AF221634 Homo sapi
3	3003	96.2	3106	6	AX342633 Sequence
4	2957	94.8	3143	6	AX354793 Sequence
5	2889	92.6	4829	6	AR631283 Sequence
6	2889	92.6	4829	6	AX608735 Sequence
7	2767	88.7	4535	8	BC040203 Homo sapi
8	2756	88.3	3030	8	AY354202 Homo sapi
9	2547	81.6	2649	6	AX354795 Sequence
10	2495	80.0	2656	8	AY172659 Homo sapi
11	2495	80.0	2671	6	AR631278 Sequence
12	2495	80.0	2671	6	AX608725 Sequence
13	2315	74.2	2797	6	AR448400 Sequence
14	2128	68.2	4523	6	AR631281 Sequence
15	2128	68.2	4523	6	AX608731 Sequence
16	2128	68.2	4676	6	AR631287 Sequence
17	2128	68.2	4676	6	AX608743 Sequence
18	2116	67.8	2842	6	AX405770 Sequence

19	2081	66.7	3125	8	BC030688	Homo sapi
20	1984	63.6	4685	6	AR631288	Sequence
21	1984	63.6	4685	6	AX608745	Sequence
22	1963	62.9	2778	8	AK000290	Homo sapi
23	1450	46.5	2161	6	BD157001	Primer fo
24	1450	46.5	2161	6	AX878058	Sequence
25	1450	46.5	2161	8	AK027826	Sequence
26	1377	44.1	2668	6	AX405771	Homo sapi
27	1191	38.2	4309	6	AR631284	Sequence
28	1191	38.2	4309	6	AX608737	Sequence
29	1164	37.3	2510	6	AX338497	Sequence
30	1023	32.8	1669	6	AR651454	Sequence
31	1023	32.8	1669	8	AF221636	Sequence
32	878	28.1	2098	6	CQ721272	Homo sapi
33	877	28.1	1356	6	AR631282	Sequence
34	877	28.1	1356	6	AX608733	Sequence
35	790	25.3	1197	6	AR651453	Sequence
36	790	25.3	1197	8	AF221635	Homo sapi
37	789	25.3	1083	6	AR651455	Sequence
38	789	25.3	1083	8	AF221637	Homo sapi
39	760	24.4	832	6	AR631286	Sequence
40	760	24.4	832	6	AX608741	Sequence
41	739	23.7	873	6	AX083130	Sequence
42	708	22.7	735	6	AX524942	Sequence
43	700	22.4	753	6	AX524935	Sequence
44	672	21.5	1002	6	AX524939	Sequence
45	501	16.1	631	6	BD149852	Primer fo

ALIGNMENTS

RESULT 1  
AR651452  
LOCUS AR651452 3120 bp DNA linear PAT 20-APR-2005  
DEFINITION Sequence 2 from patent US 6881564.  
ACCESSION AR651452  
VERSION AR651452.1 GI:62795938  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3120)  
AUTHORS Abbott, C.A., and Gorrell, M.D.  
TITLE Dipeptidyl peptidases  
JOURNAL Patent: US 6881564-A 2 19-APR-2005;  
The University of Sydney; Sydney;  
AUX;

FEATURES  
source  
1..3120  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 3120; DB 6; Length 3120;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGTGTAAAGCTCCGAGGCGCAAGCCGCTGCTACTGCGCGCGCTTCTTACTGCGG 60  
Db 1 AAGTGTAAAGCTCCGAGGCGCAAGCCGCTGCTACTGCGCGCGCTTCTTACTGCGG 60  
Qy 61 GTTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
Db 61 GTTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
Qy 121 GAGTGGAGCG 180  
Db 121 GAGTGGAGCG 180  
Qy 181 TCCGGGCG 240  
Db 181 TCCGGGCG 240

Qy	241	CTGGGTGTTGAGATATTTTGAAACTGCGGACTGTGAGGAGAAATATGAAATCAGAGATCGG	300
Db	241		
Qy	241	CTGGGTGTTGAGATATTTTGAAACTGCGGACTGTGAGGAGAAATATGAAATCAGAGATCGG	300
Db	301	CTTAAATGCGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT	360
Qy	301	CTTAAATGCGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT	360
Db	361	GCGGATACCAGAAAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT	420
Qy	361	GCGGATACCAGAAAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT	420
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTAACCTTGCCCATGCT	480
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTAACCTTGCCCATGCT	480
Db	481	GGTGAGAACAGAGAAAAATACACTGCTTTTATTCCTGAAATTCCTGAAACTATCAATAGAGCA	540
Qy	481	GGTGAGAACAGAGAAAAATACACTGCTTTTATTCCTGAAATTCCTGAAACTATCAATAGAGCA	540
Db	541	GCAGTCTTAAATGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	541	GCAGTCTTAAATGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	601	GGAAATGTAATCTTCGAGAGAGAGACTAATTAAGGAGAAAGAAACGATTTTGGACACACTGGACTAT	660
Qy	601	GGAAATGTAATCTTCGAGAGAGAGACTAATTAAGGAGAAAGAAACGATTTTGGACACACTGGACTAT	660
Db	661	ATTGCTCTTACGATATACCAAGAGAAAGTGGAACATTTCTGTTTCAAGCCGGTAGTGGA	720
Qy	661	ATTGCTCTTACGATATACCAAGAGAAAGTGGAACATTTCTGTTTCAAGCCGGTAGTGGA	720
Db	721	ATTTATCAGTAAAGATGAGGAGGCCACAAAGGATTTACGCAACAACTTTAAGGCCCAAT	780
Qy	721	ATTTATCAGTAAAGATGAGGAGGCCACAAAGGATTTACGCAACAACTTTAAGGCCCAAT	780
Db	781	CTAGTGGAAACTAGTGTGTCACCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Qy	781	CTAGTGGAAACTAGTGTGTCACCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Db	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAAACATGCTAACACAGAGAA	900
Qy	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAAACATGCTAACACAGAGAA	900
Db	901	GAAAGGAGACTCACTTATGTGCACAAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA	960
Qy	901	GAAAGGAGACTCACTTATGTGCACAAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA	960
Db	961	GCTGGAGTCGCTACCTTTGTTCTCAAGAAGAAATTTGATAGATATTTCTGGCTATTGTGG	1020
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCAAGAAGAAATTTGATAGATATTTCTGGCTATTGTGG	1020
Db	1021	TGTCCAAAGCTGAAACAACTCCAGGTGGTAAATTCCTAGAAATCTATATGAGAA	1080
Qy	1021	TGTCCAAAGCTGAAACAACTCCAGGTGGTAAATTCCTAGAAATCTATATGAGAA	1080
Db	1081	AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAACCAAGGAGG	1140
Qy	1081	AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAACCAAGGAGG	1140
Db	1141	GCGAGTATCATTCGGTATCTTAAACAGGTPACAGCAAAATCCTAAAGTCATTTTAAGATG	1200
Qy	1141	GCGAGTATCATTCGGTATCTTAAACAGGTPACAGCAAAATCCTTAAAGTCATTTTAAGATG	1200
Db	1201	TCAGAAAATAATGATTTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACTAATT	1260
Qy	1201	TCAGAAAATAATGATTTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACTAATT	1260
Db	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAAATATATTGCCAGAGCTGGATGCACTCCT	1320
Qy	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAAATATATTGCCAGAGCTGGATGCACTCCT	1320

Qy	1321	GAGGAAAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG	1380
Db	1321	GAGGAAAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG	1380
Qy	1381	TTGATCTCACCTGAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Db	1381	TTGATCTCACCTGAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Qy	1441	ATTGAGTCAGTCGCTGATTTCTGTGACGCCACTAAATTATCTATGAAGAAACAACAGACATC	1500
Db	1441	ATTGAGTCAGTCGCTGATTTCTGTGACGCCACTAAATTATCTATGAAGAAACAACAGACATC	1500
Qy	1501	TGGATAAATATCCATGACATCTTTTCATGTGTTTTTCCCACAAAGTCACGAAGAGGAAATTTGAG	1560
Db	1501	TGGATAAATATCCATGACATCTTTTCATGTGTTTTTCCCACAAAGTCACGAAGAGGAAATTTGAG	1560
Qy	1561	TTTATTTTTTGGCTCTGGAATGCAAAACAGGTTTCCGTTCATTATACAAAATTTACATCTATT	1620
Db	1561	TTTATTTTTTGGCTCTGGAATGCAAAACAGGTTTCCGTTCATTATACAAAATTTACATCTATT	1620
Qy	1621	TTAAGGGAAGCAATATTAACGATCCAGTGTGGGCTGCCTGCCAAGTGATTTTCAAG	1680
Db	1621	TTAAGGGAAGCAATATTAACGATCCAGTGTGGGCTGCCTGCCAAGTGATTTTCAAG	1680
Qy	1681	TGTCCTATCAAGAGGAGATAGCAATTAACAGTGTGTAATGGAAGTCTTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAGAGGAGATAGCAATTAACAGTGTGTAATGGAAGTCTTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Db	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Qy	1801	TCCCTTTTAGAGCATACCTGTACGTAGTCAGTTACGTAAATCTCGGAGGTTGACAAAGG	1860
Db	1801	TCCCTTTTAGAGCATACCTGTACGTAGTCAGTTACGTAAATCTCGGAGGTTGACAAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAGCACTGTGACTTCTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAGCACTGTGACTTCTTTATA	1920
Qy	1921	AGTAAAGTATAGTAAACAGAGAATCCACACTGTGTGTCCCTTTTACAAGCTATCAAAGTCT	1980
Db	1921	AGTAAAGTATAGTAAACAGAGAATCCACACTGTGTGTCCCTTTTACAAGCTATCAAAGTCT	1980
Qy	1981	GAAGATGACCCAACTTGCACAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGT	2040
Db	1981	GAAGATGACCCAACTTGCACAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTATCTCTCCAGAAAATTTTCTCTTTTGAAGAGTACTACTGGATTTACA	2100
Db	2041	CCTCTTCTGACTATCTCTCCAGAAAATTTTCTCTTTTGAAGAGTACTACTGGATTTACA	2100
Qy	2101	TTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Db	2101	TTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTTGGTCTCTCAGGTGCAAGTTGGTGAATATCGTTTAAAGNGTCAG	2220
Db	2161	CTGTTTCATATATGTTGGTCTCTCAGGTGCAAGTTGGTGAATATCGTTTAAAGNGTCAG	2220
Qy	2221	TATTTCCGTTGAATACCTTAGGCCTCTCAGGTTAATGTGGTTGTAGTAGACACACAGG	2280
Db	2221	TATTTCCGTTGAATACCTTAGGCCTCTCAGGTTAATGTGGTTGTAGTAGACACACAGG	2280
Qy	2281	GGATCTGTCAACGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2340
Db	2281	GGATCTGTCAACGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2340
Qy	2341	GAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGAC	2400
Db	2341	GAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGAC	2400
Qy	2401	TTAGATCTGTGGGCATCCACGGCTGGTCCCTATGGAGGATACCTCTCCCTGTGGCATTA	2460

FEATURES		Location/Qualifiers	
source		1..3127	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/chromosome="15"	
		/map="15q22"	
gene		1..3127	
		/gene="DPP8"	
CDS		214..2862	
		/gene="DPP8"	
		/note="serine protease; member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; alternatively spliced forms of this mRNA are found in AF221635, AF221636 and AF221637"	
		/codon_start=1	
		/product="dipeptidyl peptidase 8"	
		/protein_id="AAG29766.1"	
		/db_xref="GI:11095188"	
		/translation="MAAAETEQLGVIEFETADCBENIESQDRPKLEPFYVERYSWSQ LKLLADTRKYHGMYMAKAPHDFMFKRNDPDGPHSDRIYYLAMSGENRNTLPYSEI PKTINRAAVLMLSWKPLLDLFOATLDYGMYSREELLRRKRIGTVGLASDYDHQSGG TFLFQAGSGIYHVKGDPGQFTQOPLRPNLVETSCPNRMDPKLPADDMIAFTHSN DIWLSNITREERRLTYVHNELANMEEDARSAGVATFVLOEEDRYSGYWCWCPKAETT PSQKILIRLIYEENDESEVEI IHVTSFMLETRRADSFYPKGTGTANPKVTFMSEIMI DAEGRIIDVDKELIQPFELFEGVEYIARAGMTPEGYAMSILLDRSOTRLQIVLIS PELFIPVEDDDVMERQRLIESVDSVTPLLIYYEETDIIWINIHDI FHVFPQSHHEEIEF IFASECKTGPRHLVKITSLIKESKYKRSSGGLPAPSDPKCPIKEBIAITSGEWEVLGR HGNSIQVDEVRRLVYFEGTKDSPLEHLLYVSYVNPGEVTRLTDRGYSHSCSICQHCND FFTISKYQKNPCHVCISLYKLSSPEDDPTCKTKFETATILDSAGPLDTPPEIPFSFES TFLYGLMYLKHFDLQPGKPYTVLFYIGQPVQLVNNRFGKVKYPRNLTLASLGYV VVVIDNRGSCHRGLKFEKAGKYKMGQIEIDDOVEGLQYLAIRYDFIDLDYVGIHWSY GGYSLSLMALMORSDI FRVALAGAPVTLWIFYDTGYTERYMGHPDQNEOGYVLGVSAMQ AEPSPSPNRLLLHGHFDENVHFAHTSILLSFLVRACKPYDLQIYPPERHSIRVPES GEHYELHLHLHQENLGSRIAALKV1"	
ORIGIN			
Query Match		100.0%; Score 3120; DB 8; Length 3127;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 3120; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	AAGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
QY	61	CGTTTCGCGCGCTGGGTTGTTCACCGCGCGCGCGCGCGAGGAAGCCACTGCAACCCAGGACCG	120
Db	61	CGTTTCGCGCGCTGGGTTGTTCACCGCGCGCGCGCGCGAGGAAGCCACTGCAACCCAGGACCG	120
QY	121	GAGTGGAGGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCACTGTCGGGACGG	180
Db	121	GAGTGGAGGCGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCACTGTCGGGACGG	180
QY	181	TCCGGGCGGGCGCGGGGGGAGGAAAATGCAACATGGCAGCAGCAATGGAACACAGACAG	240
Db	181	TCCGGGCGGGCGCGGGGGGAGGAAAATGCAACATGGCAGCAGCAATGGAACACAGACAG	240
QY	241	CTGGGTGTTCAGATATTTTGAATCTCGGACTGTGAGGAGAAATATTGAATTCACAGATCGG	300
Db	241	CTGGGTGTTCAGATATTTTGAATCTCGGACTGTGAGGAGAAATATTGAATTCACAGATCGG	300
QY	301	CCTAAATTGGAGCGCTTTTATGTTGAGCGGCTATTCCTGGAGTTCAGCTTAAAAAGCTGCTT	360
Db	301	CCTAAATTGGAGCGCTTTTATGTTGAGCGGCTATTCCTGGAGTTCAGCTTAAAAAGCTGCTT	360
QY	361	CCCAGATACCAAGAAAATATCATATGCTACTATGATGGCTAAGGCCACCATATGATTTTATGTT	420
Db	361	CCCAGATACCAAGAAAATATCATATGCTACTATGATGGCTAAGGCCACCATATGATTTTATGTT	420
QY	421	GTGAAGAGGAATGATTCAGATGGACCTCATTTTACAGACAGATCTTATACCTTGGCCATGTCT	480
Db	421	GTGAAGAGGAATGATTCAGATGGACCTCATTTTACAGACAGATCTTATACCTTGGCCATGTCT	480

Qy	481	GGTGAGAAACAGAGAAATACACTGTTTATTTCTGAAATTCCTCAAAATCCCAAACTATCAATAGAGCA	540	Db	1561	TTTTATTTTGGCTCTCGAATGCAAAAACAGGTTTCGGTCATTTTATACAAAATTTACATCTATT	1520
Db	481	GGTGAGAAACAGAGAAATATACTGTTTATTTCTGAAATTCCTCAAAATCCCAAACTATCAATAGAGCA	540	Qy	1621	TTAAAGAAAGCAAAATATAAAACGATCCAGTGGTGGGCTGCTCTGCTCCAAAGTGATTTCAAG	1680
Qy	541	GCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600	Db	1621	TTAAAGAAAGCAAAATATAAAACGATCCAGTGGTGGGCTGCTCTGCTCCAAAGTGATTTCAAG	1680
Db	541	GCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600	Qy	1681	TGTCCTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTTGGCCGGCAT	1740
Qy	601	GGATGTATTCTCGAGAAGAAGAACTATTAAAGAGAAGAAAACGCAATTTGGAACAGTCGGA	660	Db	1681	TGTCCTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTTGGCCGGCAT	1740
Db	601	GGATGTATTCTCGAGAAGAAGAACTATTAAAGAGAAGAAAACGCAATTTGGAACAGTCGGA	660	Qy	1741	GGATCTAAATATCAAACTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Qy	661	ATTGCTTTCTTACGATTATCACCAAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA	720	Db	1741	GGATCTAAATATCAAACTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Db	661	ATTGCTTTCTTACGATTATCACCAAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA	720	Qy	1801	TCCCTTTAGAGCATCACCTGTACGTAGTACGTAGTACGTAAATCTCGAGAGGTCACAAAG	1860
Qy	721	ATTATCACCTAAAGATGGAGGCCACAAAGATTTTACGCAACACCTTTTAAAGCCCAAT	780	Db	1801	TCCCTTTAGAGCATCACCTGTACGTAGTACGTAGTAAATCTCGAGAGGTCACAAAG	1860
Db	721	ATTATCACCTAAAGATGGAGGCCACAAAGATTTTACGCAACACCTTTTAAAGCCCAAT	780	Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCACTGACGTGTGACTTCTTTATA	1920
Qy	781	CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCCGCTGATCCA	840	Db	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCACTGACGTGTGACTTCTTTATA	1920
Db	781	CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCCGCTGATCCA	840	Qy	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCCCTTTACAAAGCTATCAAGTCT	1980
Qy	841	GACTGGATTTGCTTTTATACATAGCAACGATATTTTGGATATCTAACATCGTAAACAGAGAA	900	Db	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCCCTTTACAAAGCTATCAAGTCT	1980
Db	841	GACTGGATTTGCTTTTATACATAGCAACGATATTTTGGATATCTAACATCGTAAACAGAGAA	900	Qy	1981	GAAGATGACCCAACTTGGCAAAAACAAAGGAAATTTTGGGCCACATTTTGGATTCAGAGGT	2040
Qy	901	GAAGGAGACTCATTTATGTGCACAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA	960	Db	1981	GAAGATGACCCAACTTGGCAAAAACAAAGGAAATTTTGGGCCACATTTTGGATTCAGAGGT	2040
Db	901	GAAGGAGACTCATTTATGTGCACAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA	960	Qy	2041	CTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGATGTACTACCTGGATTTACA	2100
Qy	961	GCTGGAGTCGCTACCTTTGTTCTTCCAAAGAAGAAATTTGATAGATATTTCTGGCTATTTGGTG	1020	Db	2041	CTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGATGTACTACCTGGATTTACA	2100
Db	961	GCTGGAGTCGCTACCTTTGTTCTTCCAAAGAAGAAATTTGATAGATATTTCTGGCTATTTGGTG	1020	Qy	2101	TTGTATGGGATGCTCTCAAGGCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTG	2160
Qy	1021	TGTTCAAAAGCTGAAACAACTCCACGTGGTGTGTAATAATTTAGAAATTTCTATATGAAGAA	1080	Db	2101	TTGTATGGGATGCTCTCAAGGCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTG	2160
Db	1021	TGTTCAAAAGCTGAAACAACTCCACGTGGTGTGTAATAATTTAGAAATTTCTATATGAAGAA	1080	Qy	2161	CTGTTTCATATATGGTGGTCTCTCAGGTGCAATTTGGTGAATAATCGGTTTAAAGAGGTCAAG	2220
Qy	1081	AATCATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTGGAAAACAGGAGG	1140	Db	2161	CTGTTTCATATATGGTGGTCTCTCAGGTGCAATTTGGTGAATAATCGGTTTAAAGAGGTCAAG	2220
Db	1081	AATCATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTGGAAAACAGGAGG	1140	Qy	2221	TATTTCCGCTTGAATACCTCTAGCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGG	2280
Qy	1141	GCAGATTCATTCGGTTATCCTTAAACAGGTACAGCAAACTCTTAAAGTCACCTTTTAAAGATG	1200	Db	2221	TATTTCCGCTTGAATACCTCTAGCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAGG	2280
Db	1141	GCAGATTCATTCGGTTATCCTTAAACAGGTACAGCAAACTCTTAAAGTCACCTTTTAAAGATG	1200	Qy	2281	GGATCCTGTCAACCGAGGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAATA	2340
Qy	1201	TCAGAAATATGATGATGCTGAAGAGGATCATAGATGTCATAGATTAAGTCACTTTTAAAGT	1260	Db	2281	GGATCCTGTCAACCGAGGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAATA	2340
Db	1201	TCAGAAATATGATGATGCTGAAGAGGATCATAGATGTCATAGATTAAGTCACTTTTAAAGT	1260	Qy	2341	GAATTTGACCAATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	2400
Qy	1261	CAACCTTTTGAGATCTTATTTGAAGGATGAAATATATTTGCCAGAGCTGGATGACTCCT	1320	Db	2341	GAATTTGACCAATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	2400
Db	1261	CAACCTTTTGAGATCTTATTTGAAGGATGAAATATATTTGCCAGAGCTGGATGACTCCT	1320	Qy	2401	TTAGATCGTGGGATCCACGGCTGGTCTATGGAGGATACCTCTCCCTGATGGCATT	2460
Qy	1321	GAGGAAAATATGCTTGGTCCATCTCTAGATGCGTCCAGATCCAGATCAGATAGTG	1380	Db	2401	TTAGATCGTGGGATCCACGGCTGGTCTATGGAGGATACCTCTCCCTGATGGCATT	2460
Db	1321	GAGGAAAATATGCTTGGTCCATCTCTAGATGCGTCCAGATCCAGATCAGATAGTG	1380	Qy	2461	ATGAGAGGTGAGATATCTTTCAGGGTGTCTATTTGCTGGGGCCCGCTGCTCTGGATC	2520
Qy	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440	Db	2461	ATGAGAGGTGAGATATCTTTCAGGGTGTCTATTTGCTGGGGCCCGCTGCTCTGGATC	2520
Db	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440	Qy	2521	TTCTATGATACAGGATACAGGAAACGTTATATGGGTCAACCTGACCAAGATGAACAGGGC	2580
Qy	1441	ATTGAGTCAGTCCCTGATTTCTGTGAGCCCACTAAATTTCTATGAAGAAACCAACAGATC	1500	Db	2521	TTCTATGATACAGGATACAGGAAACGTTATATGGGTCAACCTGACCAAGATGAACAGGGC	2580
Db	1441	ATTGAGTCAGTCCCTGATTTCTGTGAGCCCACTAAATTTCTATGAAGAAACCAACAGATC	1500	Qy	2581	TATTTACTTAGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTCGAAACCAAAATCGTTTA	2640
Qy	1501	TGGATAAATATCCATGACATCTTTTCTATGTTTTTCCCAAGTTCAGGAAGGAAATTTGAG	1560	Db	2581	TATTTACTTAGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTCGAAACCAAAATCGTTTA	2640
Db	1501	TGGATAAATATCCATGACATCTTTTCTATGTTTTTCCCAAGTTCAGGAAGGAAATTTGAG	1560	Qy	2641	CTGCTCTTACATGTTTCTCTGGATGAGAAATGTCATTTTGGACATACACAGTATATTACTG	2700
Qy	1561	TTTTATTTTGGCTCTGAAATCAAAAACAGGTTTTCGGTCAATTTTATACAAAATTTACATCTATT	1620				

Db	2641	CTGCTCTTACATGGTTTCTCTGGATGAGATGTCCATTTTGGCATACTACAGTATATTACTG	2700
Qy	2701	AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGACAC	2760
Db	2701		2760
Qy	2761	AGCATAAGAGTTCTGTAATCGGAGAGAACATTTATGAATCGTCTTTTGCACTACCTTCAA	2820
Db	2761	AGCATRAAGTGTTCTGTAATCGGAGAGAACATTTATGAATCGTCTTTTGCACTACCTTCAA	2820
Qy	2821	GAACACCTTGGATCAGCTATTTGCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAAC	2880
Db	2821	GAACACCTTGGATCAGCTATTTGCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAAC	2880
Qy	2881	TCTCTGGTATACATCTGGCTATTTAAACCAATGAGGAGGTTTAATCAACAGAGAAAACACAGA	2940
Db	2881	TCTCTGGTATACATCTGGCTATTTAAACCAATGAGGAGGTTTAATCAACAGAGAAAACACAGA	2940
Qy	2941	ATTGATCATCATTTTGTATACCTGCCATGTAATCTACTCTGAAATAAATGTGGTG	3000
Db	2941	ATTGATCATCATTTTGTATACCTGCCATGTAATCTACTCTGAAATAAATGTGGTG	3000
Qy	3001	CCATGACGGGCTCTACGGTTTGGTAGTAATCTTAATACCTTAACCCACATGCTCAAAA	3060
Db	3001	CCATGACGGGCTCTACGGTTTGGTAGTAATCTTAATACCTTAACCCACATGCTCAAAA	3060
Qy	3061	TCAATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAA	3120
Db	3061	TCAATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAA	3120
RESULT 3			
AX342633			
LOCUS			
Sequence 30 from Patent WO0198468.			
DEFINITION			
ACCESSION			
VERSION			
AX342633.1 GI:18152030			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
AUTHORS			
Yue, H., Elliott, V.S., Gandhi, A.R., Lal, P., Au-Young, J.,			
Tribouley, C.M., Deleage, A.M., Baughn, M.R., Nguyen, D.B., Lee, E.A.,			
Hafalia, A., Khan, F.A., Wallia, N.K., Yao, M.G., Lu, D.A., Patterson, C.,			
Tang, Y.T., Walsh, R.T., Azimzai, Y., Runkumar, J., Xu, Y. and Reddy, R.			
Patent: WO 0198468-A 30 27-DEC-2001;			
JOURNAL			
Incyte Genomics, Inc. (US)			
FEATURES			
source			
1. 3106			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
/note="Incyte ID No: 7160544CB1"			
ORIGIN			
Query Match 96.2%; Score 3003; DB 6; Length 3106;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 3103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	13	CTCCAGGCGCAGGCGCTGCTACTGCGCGCTGCTTCTTAGTGCGCGCTTCGCGCT	72
Db	2	CTCCAGGCGCAGGCGCTGCTACTGCGCGCTGCTTCTTAGTGCGCGCTTCGCGCT	61
Qy	73	GGGTTGTACCGGCGCGCGCGCGGAGCACTCAACAGGACCCGAGTGGAGCGG	132
Db	62	GGGTTGTACCGGCGCGCGCGCGGAGCACTCAACAGGACCCGAGTGGAGCGG	121
Qy	133	CGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGACCTCGGAGCGGTCGCGCGGCG	192
Db	122	CGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGACCTCGGAGCGGTCGCGCGGCG	181



QY	1273	ATTCTATTTGAAGAGTTGAATATATATATGCGAGAGCTGGATGGA	1332	Db	2342	CAGGTGGAAAGGACTCCAATATCTAGCTTCTCGATATGATTTT	2400			
Db	1262	ATTCTATTTGAAGAGTTGAATATATATGCGAGAGCTGGATGGA	1321	QY	2413	GGCATCCACGGCTGGTCTTATGAGGAGTACCTCTCCCTGATGG	2472			
QY	1333	GCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCCTACAGAT	1392	Db	2402	GGCATCCACGGCTGGTCTTATGAGGAGTACCTCTCCCTGATGG	2461			
Db	1322	GCTTGGTCCATCTCTACTAGATCGCTCCAGAGCTCGCCTACAG	1381	QY	2473	GATATCTTTCAGGGTTGCTATTTGCTGGGGCCCGAGTCACTCT	2532			
QY	1393	GAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAG	1452	Db	2462	GATATCTTTCAGGGTTGCTATTTGCTGGGGCCCGAGTCACTCT	2521			
Db	1382	GAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAG	1441	QY	2533	GGATACACGGAACGTTTATATGGGTCACTGACCAAGATGAAC	2592			
QY	1453	CCTGATCTGTGAGCGCCACTAAATATCTATGAAGAAACAAAG	1512	Db	2522	GGATACACGGAACGTTTATATGGGTCACTGACCAAGATGAAC	2581			
Db	1442	CCTGATCTGTGAGCGCCACTAAATATCTATGAAGAAACAAAG	1501	QY	2593	TCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAA	2652			
QY	1513	CATGACATCTTTCAATGTTTTTCCCAAGTTCAGAGAGGAAAT	1572	Db	2582	TCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAA	2641			
Db	1502	CATGACATCTTTCAATGTTTTTCCCAAGTTCAGAGAGGAAAT	1561	QY	2653	GGTTTCTCTGATGAGAAATGTCCTATTTTGCACATATATTA	2712			
QY	1573	TCTGAAATGCRAACAGGTTTCGGTCATTTATACAAATTTACA	1632	Db	2642	GGTTTCTCTGATGAGAAATGTCCTATTTTGCACATATATTA	2701			
Db	1562	TCTGAAATGCRAACAGGTTTCGGTCATTTATACAAATTTACA	1621	QY	2713	AGGCTCTGGAAGCCATATGATTTTACAGATCTATCTCCTCA	2772			
QY	1633	AAATATAAACATCCAGTGTGGGCTGCTGCTCCAGTGATTTCA	1692	Db	2702	AGGCTCTGGAAGCCATATGATTTTACAGATCTATCTCCTCA	2761			
Db	1622	AAATATAAACATCCAGTGTGGGCTGCTGCTCCAGTGATTTCA	1681	QY	2773	CCTGAATCGGGAGAACATTTATGAACCTGCATCTTTTGGAC	2832			
QY	1693	GAGGAGATAGCAATATACAGTGTGGAATGGGAAGTTCTTGG	1752	Db	2762	CCTGAATCGGGAGAACATTTATGAACCTGCATCTTTTGGAC	2821			
Db	1682	GAGGAGATAGCAATATACAGTGTGGAATGGGAAGTTCTTGG	1741	QY	2833	TCACGTATTTGCTGCTCTAAAAGTGATATAATTTTGAACCT	2892			
QY	1753	CAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCCA	1812	Db	2822	TCACGTATTTGCTGCTCTAAAAGTGATATAATTTTGAACCT	2881			
Db	1742	CAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCCA	1801	QY	2893	ACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAA	2952			
QY	1813	CATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGT	1872	Db	2882	ACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAA	2941			
Db	1802	CATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGT	1861	QY	2953	ATTTTGATACCTGCCATGTAAACATCTACTCTCGAAAAATAA	3012			
QY	1873	GGCTACTCACATCTTCTGCTGCATCAGTCAGCACTGTGACT	1932	Db	2942	ATTTTGATACCTGCCATGTAAACATCTACTCTCGAAAAATAA	3001			
Db	1862	GGCTACTCACATCTTCTGCTGCATCAGTCAGCACTGTGACT	1921	QY	3013	CTACGGTTTGTGGTAGTAAATCTAATACCTTAACCCCAATG	3072			
QY	1933	AACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAG	1992	Db	3002	CTACGGTTTGTGGTAGTAAATCTAATACCTTAACCCCAATG	3061			
Db	1922	AACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAG	1981	QY	3073	TATTCCTGAGAGACCCAGCAATACCATTAAGAATTTACTAAAA	3117			
QY	1993	ACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTC	2052	Db	3062	TATTCCTGAGAGACCCAGCAATACCATTAAGAATTTACTAAAA	3106			
Db	1982	ACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTC	2041	RESULT 4						
QY	2053	TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATT	2112	LOCUS	AX354793	3143 bp	DNA linear PAT 06-FEB-2002			
Db	2042	TATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATT	2101	DEFINITION	Sequence 1 from Patent WO0179473.					
QY	2113	CTCTAAGAGCTCATGATCTACAGCTCGGAAGAAATATCTCTA	2172	ACCESSION	AX354793					
Db	2102	CTCTAAGAGCTCATGATCTACAGCTCGGAAGAAATATCTCTA	2161	VERSION	AX354793.1	GI:18619526				
QY	2173	GGTGGTCTCAGGTGCAAGTTGGTGAATATTCGGTTTAAAGGAG	2232	KEYWORDS	Homo sapiens (human)					
Db	2162	GGTGGTCTCAGGTGCAAGTTGGTGAATATTCGGTTTAAAGGAG	2221	SOURCE	Homo sapiens					
QY	2233	AATACCTTAGCCTCTTAGTTATGTGGTTGTAGTGATAGACAA	2292	ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. 1					
Db	2222	AATACCTTAGCCTCTTAGTTATGTGGTTGTAGTGATAGACAA	2281	REFERENCE				Meyers, R.A. and Williamson, M.		
QY	2293	CGAGGGCTTAAATTTGAAGCGCGCTTTTAAATATAAATGGGT	2352	AUTHORS				21953, a human prolyl oligopeptidase family member and uses thereof		
Db	2282	CGAGGGCTTAAATTTGAAGCGCGCTTTTAAATATAAATGGGT	2341	TITLE				Patent: WO 0179473-A 1 25-OCT-2001;		
QY	2353	CAGGTGGAAAGGACTCCAATATCTAGCTTCTCGATATGATTT	2412	JOURNAL				Millennium Pharmaceuticals, Inc. (US)		
Db	2342	CAGGTGGAAAGGACTCCAATATCTAGCTTCTCGATATGATTT	2400	FEATURES	Location/Qualifiers					
				source	1. 3143					
					/organism="Homo sapiens"					
					/mol_type="unassigned DNA"					

CDS	/db_xref="taxon:9606" 229..2877		/note="unnamed protein product"		/codon_start=1		/protein_id="CAD22990.1"		/db_xref="GI:18619527"		/translation="MAAAMETEQLGVIEFETADCEBIESQDRPKLEPFYVERYSWSQ LKKLLADPRKHYGYMAKAPHDFMFKVNDPDGPHSDRIYVYLAWSGENRNTLFSYSEI PKTINRAAVLMWSKPLDLFQATLDYGYMSREELLRRKRIGTVGTAASYDYHQGG TFLFAGSGSIVHVKDGGPQGTQOPLRPNLVETSCNIRMDPKLPADPDWIAPIHSH DIWISNVTREERLTYVHNELANMEEDARSAGVATFVLQBEFDRYSGYMKPAET PSGGKILRILENDESEVEI IHVTSMPLETRADSFRPKGTANPKVTFKMEIMI DABGRIIDIVIKELIOPFEILFSGVEIARAGWTEGKYAWSIILLDRSOTRLOIIVLS PELFPVEDDWMERQLIESVPDSVTFPLIIEYETDINIHDIHFVPPQSHSEIEF IFASCKTGFRLHYKITSILKESIKRSGSGLPAPSDFKPIKEBIAITSGEWEVLGR HGSNVDVFRVLYFEGTKDPSLEHLHYVSYVNPGEVTRITDRGYSHSCCISQHCD FFISKYSNOKPHCVSLYKSSPEDDPTCKTFEATILDSAGPLDVTPEIEFSFVS TTGFTLXGMLYKPHDLOPKKYPVLPIYGGQVOLVNNRFGVKYFRLNLTASLGTV VVVDNRSGCHRGGLKFGKAFKYMKGQIEIDQVEGLQYLASRYDFIDLDRLDGHNSY GGYLSLMALMQRSDI FRVAIAGAPVTLMIFYDTGYTERYMGHPDQNEGGYLYLGSVAMQ AEKFPSEPNRLLHGLFDENVHFAHTSILLSPFVRAGKPYDLQIYPOERHSIRVPS GEHYELHLLHYLQENLGSRTAALKVI"		ORIGIN	Query Match Best Local Similarity Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	94.8%; Score 2957; DB 6; Length 3143; 99.9%; Pred. No. 0;	11	GCTCCGAGCCCAAGCCGCTGCTACTGCCGCGCTGCTTCTTAGTGGCGGCTTCGCCGC	70	11	GCTCCGAGCCCAAGCCGCTGCTACTGCCGCGCTGCTTCTTAGTGGCGGCTTCGCCGC	85	71	CTGGGTGTGTCACCGCGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGAGGC	130	86	CTGGGTGTGTCACCGCGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGAGGC	145	131	GCGCGAGCATGAAGCGCGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGAGGC	190	146	GCGCGAGCATGAAGCGCGCGCGCGCGGAGGAGCCACTGCAACAGGACCGGAGTGAGGC	205	191	GCGCGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTG	250	206	GCGCGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTG	265	251	AGATATTTGAACTCGGACCTGTGAGGAGAAATTTGAATCAGAGGATCGGCTAAATTTGG	310	266	AGATATTTGAACTCGGACCTGTGAGGAGAAATTTGAATCAGAGGATCGGCTAAATTTGG	325	311	AGCTTTTATGTTGAGCGGTATCTCTGAGTACGCTTAAAGAGCTGTGCGGATACCA	370	326	AGCTTTTATGTTGAGCGGTATCTCTGAGTACGCTTAAAGAGCTGTGCGGATACCA	385	371	GAAATATCATGCTACATGATGGCTAAGGACCAACATGATTTTATGTTGTGAAGAGA	430	386	GAAATATCATGCTACATGATGGCTAAGGACCAACATGATTTTATGTTGTGAAGAGA	445	431	ATGATCCAGATGGAACCTCATTTGAGAGAAATCTAATTTACCTTGGCATGTCTGGTGAGACA	490	446	ATGATCCAGATGGAACCTCATTCAGACAGATCTATTACCTTGCCATGTCTGGTGAGACA	505	491	GAGAAATACATGTTTATTTCTGAAATTTCCAAACTATCAATAGAGCAGAGCTTTAA	550	506	GAGAAATACATGTTTATTTCTGAAATTTCCAAACTATCAATAGAGCAGAGCTTTAA	565	551	TGCTCTTTGAGAGCTCTTTTGGATCTTTTTCAGGCAACATGGAATGGAATGATTT	610	566	TGCTCTTTGAGAGCTCTTTTGGATCTTTTTCAGGCAACATGGAATGGAATGATTT	625	611	CTCGAGAGAGAACTATTTAAG
-----	------------------------------------	--	---------------------------------	--	----------------	--	--------------------------	--	------------------------	--	---	--	--------	---	--	----	---	----	----	---	----	----	---	-----	----	---	-----	-----	--	-----	-----	--	-----	-----	---	-----	-----	---	-----	-----	---	-----	-----	---	-----	-----	---	-----	-----	---	-----	-----	--	-----	-----	--	-----	-----	--	-----	-----	---	-----	-----	--	-----	-----	--	-----	-----	---	-----	-----	---	-----	-----	---

Db 1766 TCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCAACAAAGACTCCCCCTTTAG 1825  
Qy 1811 AGCATCACCTGTACGTAGTACAGTTACGTTAAATCTCTGGAGAGGTGCAAGGCTCACTGACC 1870  
Db 1826 AGCATCACCTGTACGTAGTACAGTTACGTTAAATCTCTGGAGAGGTGCAAGGCTCACTGACC 1885  
Qy 1871 GTGCTACTCACATTTCTTGCTGATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATA 1930  
Db 1886 GTGCTACTCACATTTCTTGCTGATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATA 1945  
Qy 1931 GTRACCAAGAAGATCACACCTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACC 1990  
Db 1946 GTRACCAAGAAGATCACACCTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACC 2005  
Qy 1991 CAATCTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGAGGCTCTCTCTG 2050  
Db 2006 CAATCTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGAGGCTCTCTCTG 2065  
Qy 2051 ACTATACTCTCCAGAAATTTCTCTTTTGAAGTACTTGGATTTACATTTGATGGGA 2110  
Db 2066 ACTATACTCTCCAGAAATTTCTCTTTTGAAGTACTTGGATTTACATTTGATGGGA 2125  
Qy 2111 TGCTCTACAAGCTCATGATCTACAGCTCTGGAAGAAATATCTTACTGTCTGTCATAT 2170  
Db 2126 TGCTCTACAAGCTCATGATCTACAGCTCTGGAAGAAATATCTTACTGTCTGTCATAT 2185  
Qy 2171 ATGTGTCTCTCAGGTGCTAGTGTGGTGAATTAATCGGTTTAAAGGATCAAGTATTCGGCT 2230  
Db 2186 ATGTGTGTCTCAGGTGCTAGTGTGGTGAATTAATCGGTTTAAAGGATCAAGTATTCGGCT 2245  
Qy 2231 TGAATACCTTAGCTCTCTAGGTTATGTGGTTGTAGTATGATAGACAAACAGGGGATCTCTGTC 2290  
Db 2246 TGAATACCTTAGCTCTCTAGGTTATGTGGTTGTAGTATGATAGACAAACAGGGGATCTCTGTC 2305  
Qy 2291 ACCGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAATGAATGAAG 2350  
Db 2306 ACCGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAATGAATGAAG 2365  
Qy 2351 ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTG 2410  
Db 2366 ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTG 2425  
Qy 2411 TGGGATCTCACGGCTGTGCTCTATGAGGATACCTCTCCCTGATGCGATTAATGAGAGGT 2470  
Db 2426 TGGGATCTCACGGCTGTGCTCTATGAGGATACCTCTCCCTGATGCGATTAATGAGAGGT 2485  
Qy 2471 CAGATATCTTCAGGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATA 2530  
Db 2486 CAGATATCTTCAGGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATA 2545  
Qy 2531 CAGGATACAGGACGTTATATGCGTCACTGACCCAGAAAGTGAACAGGGCTATTACTTAG 2590  
Db 2546 CAGGATACAGGACGTTATATGCGTCACTGACCCAGCAATGAACAGGGCTATTACTTAG 2605  
Qy 2591 GATCTGTGGGCTGCAAGCAAGAAAGTTCGCTCTGAAACCAATCGTTTACTGCTTTAC 2650  
Db 2606 GATCTGTGGGCTGCAAGCAAGAAAGTTCGCTCTGAAACCAATCGTTTACTGCTTTAC 2665  
Qy 2651 ATGGTTTCTCGGATGAGAAATGTCATTTTGCATACAGTATATTTACTGAGTTTATAG 2710  
Db 2666 ATGGTTTCTCGGATGAGAAATGTCATTTTGCATACAGTATATTTACTGAGTTTATAG 2725  
Qy 2711 TGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAGAG 2770  
Db 2726 TGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGAGAGACACAGCATAGAG 2785  
Qy 2771 TTCTCGAATCGGGAACATATATGAATCTGTTTGTGATCTTCTTCAAGAAACCTTG 2830  
Db 2786 TTCTCGAATCGGGAACATATATGAATCTGTTTGTGATCTTCTTCAAGAAACCTTG 2845  
Qy 2831 GATCAGTATTTGCTGCTCTAAAAGTGATATAATTTTGAACCTGTGTAGAACCTCTCTGGTAT 2890  
Db 2846 GATCAGTATTTGCTGCTCTAAAAGTGATATAATTTTGAACCTGTGTAGAACCTCTCTGGTAT 2905

Qy 2891 ACATCGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATC 2950  
Db 2906 ACATCGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATC 2965  
Qy 2951 ACATTTTGATACCTGCGCATGTAACATCTACTCTCTGAAAAATAAATGTGGTCCCATGCGGG 3010  
Db 2966 ACATTTTGATACCTGCGCATGTAACATCTACTCTCTGAAAAATAAATGTGGTCCCATGCGGG 3025  
Qy 3011 GTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATA 3070  
Db 3026 GTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATA 3085  
Qy 3071 CATATTCTCTGAGAGACCCAGCAATACCAATAAGAATTTACTAAAAAATAAAAA 3120  
Db 3086 CATATTCTCTGAGAGACCCAGCAATACCAATAAGAATTTACTAAAAAATAAAAA 3135

RESULT 5  
AR631283  
LOCUS AR631283 4829 bp DNA linear PAT 14-FEB-2005  
DEFINITION Sequence 12 from patent US 6844180.  
ACCESSION AR631283  
VERSION AR631283.1 GI:59770927  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4829)  
AUTHORS Oi, S., Akimanya, K.O., Riviere, P.J.M. and Junien, J.-L.  
TITLE Serine protease genes, related to DPPIV  
JOURNAL Patent: US 6844180-A 12 18-JAN-2005;  
Ferring BV;;  
NLX;

FEATURES  
source Location/Qualifiers  
1..4829  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 92.6%; Score 2889; DB 6; Length 4829;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;  
Qy 1 AAGTGCTAAAGCTTCGAGGCCAAGGCCGCTGTCTACTCGCGCGCTGCTTCTTGTAGTCCG 60  
Db 1 AAGTGCTAAAGCTTCGAGGCCAAGGCCGCTGTCTACTCGCGCGCTGCTTCTTGTAGTCCG 60  
Qy 61 CGTTTCGCGCTCGGTTGTACCGCGCGCGCGCGCGAGGAAGCCACTGCAACCAAGGACCG 120  
Db 61 CGTTTCGCGCTCGGTTGTACCGCGCGCGCGCGCGAGGAAGCCACTGCAACCAAGGACCG 120  
Qy 121 GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCACGTGCGGACGG 180  
Db 121 GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCTCCATAGCGCACGTGCGGACGG 180  
Qy 181 TCCGCGCGCGCGCGCGGGAAGAAAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240  
Db 181 TCCGCGCGCGCGCGCGGGAAGAAAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240  
Qy 241 CTGGTCTTCAGATATTTTGAATCTGCGGACTGTGAGGAGAAATTTGAAATCAAGGATCGG 300  
Db 241 CTGGTCTTCAGATATTTTGAATCTGCGGACTGTGAGGAGAAATTTGAAATCAAGGATCGG 300  
Qy 301 CCTAAATGAGGCTTTTATGTTAGCGGTATTCCTGAGTCAAGTCAAGGATCGGTT 360  
Db 301 CCTAAATGAGGCTTTTATGTTAGCGGTATTCCTGAGTCAAGTCAAGGATCGGTT 360  
Qy 361 GCGGATACCAAGAAATATCATGCTGATGATGGTAAAGGACACATGATTTTATGTTT 420  
Db 361 GCGGATACCAAGAAATATCATGCTGATGATGGTAAAGGACACATGATTTTATGTTT 420  
Qy 421 GTGAAGAGGAATGATCCAGATGGACCTCATTACAGACAGAACTATTACCTTGCCATGCT 480

|||||  
421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAATCTATTACCTTTGCCATGTCT 480  
QY  
481 GGTGAGAACAGAGAAAATACACTGTTTTTATTCTCGAAATCCCAAACTATCAATAGAGCA 540  
Db  
481 GGTGAGAACAGAGAAAATACACTGTTTTTATTCTCGAAATCCCAAACTATCAATAGAGCA 540  
QY  
541 GCAGTCTTAATGCTCTCTTGGAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
Db  
541 GCAGTCTTAATGCTCTCTTGGAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
QY  
601 GGAATGATTTCTCGAGAAGAGAACTATTAAAGAGAAAGAAAACGCAATTTGGAACAGTCGGA 660  
Db  
601 GGAATGATTTCTCGAGAAGAGAACTATTAAAGAGAAAGAAAACGCAATTTGGAACAGTCGGA 660  
QY  
661 ATTGCTTTCTACGATTTACCAAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA 720  
Db  
661 ATTGCTTTCTACGATTTACCAAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA 720  
QY  
721 ATTTATCACGTAAAAGATGAGGGCCACAGGATTTACGCAACACTTTAAGGCCCAAT 780  
Db  
721 ATTTATCACGTAAAAGATGAGGGCCACAGGATTTACGCAACACTTTAAGGCCCAAT 780  
QY  
781 CTAGTGGAACTAGTTGTCCTCCACATACCGATGGATCCAAAATTTATGCCCGCTGATCCA 840  
Db  
781 CTAGTGGAACTAGTTGTCCTCCACATACCGATGGATCCAAAATTTATGCCCGCTGATCCA 840  
QY  
841 GACTGGATTTGCTTTTATACATAGCAACGATATTGATATCTAAATCTGTAACAGAGAA 900  
Db  
841 GACTGGATTTGCTTTTATACATAGCAACGATATTGATATCTAAATCTGTAACAGAGAA 900  
QY  
901 GAAAGGAGACTCACTTATGTGCACAAATGAGCTAGCCAAACATGGAAGAAAGATGCCAGATCA 960  
Db  
901 GAAAGGAGACTCACTTATGTGCACAAATGAGCTAGCCAAACATGGAAGAAAGATGCCAGATCA 960  
QY  
961 GCTGGAGTCGCTACCTTTGTTCTCCAAGAAAGATTTGATAGATATTCCTGCTATTGGTGG 1020  
Db  
961 GCTGGAGTCGCTACCTTTGTTCTCCAAGAAAGATTTGATAGATATTCCTGCTATTGGTGG 1020  
QY  
1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTGTAATAATCTTAGAATTTCTATATGAAGAA 1080  
Db  
1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTGTAATAATCTTAGAATTTCTATATGAAGAA 1080  
QY  
1081 AATGATGAATCTGAGGTGGAATTTATCATGTTTACATCCCTATGCTGGAACAGAGG 1140  
Db  
1081 AATGATGAATCTGAGGTGGAATTTATCATGTTTACATCCCTATGCTGGAACAGAGG 1140  
QY  
1141 GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACATTTTAAAGATG 1200  
Db  
1141 GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACATTTTAAAGATG 1200  
QY  
1201 TCAGAAATTAATGATTGATGCTGGAAGAGGATCATAGATGTCATAGATAAGAACTAAT 1260  
Db  
1201 TCAGAAATTAATGATTGATGCTGGAAGAGGATCATAGATGTCATAGATAAGAACTAAT 1260  
QY  
1261 CAACCTTTTGAGATTTCTATTGGAAGGATTTGAATATATTGCCAGAGCTGGATGACTCCT 1320  
Db  
1261 CAACCTTTTGAGATTTCTATTGGAAGGATTTGAATATATTGCCAGAGCTGGATGACTCCT 1320  
QY  
1321 GAGGAAAAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGT 1380  
Db  
1321 GAGGAAAAATATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGT 1380  
QY  
1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGCGAGACTC 1440  
Db  
1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGCGAGACTC 1440  
QY  
1441 ATTGAGTCAGTCTGATCTGAGCCCACTAAATTTCTATGAGAAACACAGACATC 1500  
Db  
1441 ATTGAGTCAGTCTGATCTGAGCCCACTAAATTTCTATGAGAAACACAGACATC 1500  
QY  
1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAGTCACGAAGAGAAATTGAG 1560  
Db  
|||||

1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAGTCACGAAGAGAAATTGAG 1560  
QY  
1561 TTTATTTTTGCTCTGGAATCAAAAACAGTTTTCCGTCAATTTATACAAAATTTACATCTATT 1620  
Db  
1561 TTTATTTTTGCTCTGGAATCAAAAACAGTTTTCCGTCAATTTATACAAAATTTACATCTATT 1620  
QY  
1621 TTAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCCCTGCTCCAAAGTGATTTCAAG 1680  
Db  
1621 TTAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCCCTGCTCCAAAGTGATTTCAAG 1680  
QY  
1681 TGTCTTATCAAGAGGAGATAGCAATTAACGATGGTGAATGGAAGTCTTTGGCCGGCAT 1740  
Db  
1681 TGTCTTATCAAGAGGAGATAGCAATTAACGATGGTGAATGGAAGTCTTTGGCCGGCAT 1740  
QY  
1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCCAAGAC 1800  
Db  
1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCCAAGAC 1800  
QY  
1801 TCCCTTTTAGAGCATCACCTGTAGTACGTAGTCAGTTACGTAAATCCCTGGAGAGTGACAAGG 1860  
Db  
1801 TCCCTTTTAGAGCATCACCTGTAGTACGTAGTCAGTTACGTAAATCCCTGGAGAGTGACAAGG 1860  
QY  
1861 CTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCACTGTGACTTCTTTATA 1920  
Db  
1861 CTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCACTGTGACTTCTTTATA 1920  
QY  
1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCT 1980  
Db  
1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCT 1980  
QY  
1981 GAAATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTGGATTCAGCAGGT 2040  
Db  
1981 GAAATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACCAATTTGGATTCAGCAGGT 2040  
QY  
2041 CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100  
Db  
2041 CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100  
QY  
2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTG 2160  
Db  
2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTG 2160  
QY  
2161 CTGTTTCATATATATGTTGG- -TCCTCAGGTG CAGTTGGTGAATTAATCGGTTTAAAGAGTCA 2218  
Db  
2161 CTGTTTCATATATATGTTGGTCTCTCCTCAGGTG CAGTTGGTGAATTAATCGGTTTAAAGAGTCA 2220  
QY  
2219 AGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTTATGTTGGTTGTAGTATAGTATCA 2278  
Db  
2221 AGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTTATGTTGGTTGTAGTATAGTATCA 2280  
QY  
2279 GGGGATCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAA 2338  
Db  
2281 GGGGATCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAA 2340  
QY  
2339 TAGAAATTTAGGATCAGGTGGAAGGACTCAATATCTAGCTCTCTCGATATGATTTCAATG 2398  
Db  
2341 TAGAAATTTAGGATCAGGTGGAAGGACTCAATATCTAGCTCTCTCGATATGATTTCAATG 2400  
QY  
2399 ACTTAGATCTGTGGGATCCAGGCTGTCTCTATGGAGATACCTCTCCCTGATGGCAT 2458  
Db  
2401 ACTTAGATCTGTGGGATCCAGGCTGTCTCTATGGAGATACCTCTCCCTGATGGCAT 2460  
QY  
2459 TAATGCGAGGTCAGATATCTTCAGGTTGCTATTTGCTGGGCCCCAGTCACCTCTGTGGA 2518  
Db  
2461 TAATGCGAGGTCAGATATCTTCAGGTTGCTATTTGCTGGGCCCCAGTCACCTCTGTGGA 2520  
QY  
2519 TCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACAGATGAAACAGG 2578  
Db  
2521 TCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACAGATGAAACAGG 2580  
QY  
2579 GCTATTTATTTAGGATCTGTGGCCATGCAAGAGAAAGTTTCCCTCTGAAACCAATCGTT 2638  
Db  
2581 GCTATTTATTTAGGATCTGTGGCCATGCAAGAGAAAGTTTCCCTCTGAAACCAATCGTT 2640

```
QY 2639 TACTGCTCTTACATGGTTCTCTCGATGAGAAATGTCATTTTCCACATACACAGTATATTAC 2698
Db 2641 TACTGCTCTTACATGGTTCTCTCGATGAGAAATGTCATTTTCCACATACACAGTATATTAC 2700
QY 2699 TGAGTCTTTTATAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAC 2758
Db 2701 TGAGTCTTTTATAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAC 2760
QY 2759 ACAGCATAGAAGTTCTGAAATCGGAGAACATATGAACTGATCTTTTGACCTACCTTC 2818
Db 2761 ACAGCATAGAAGTTCTGAAATCGGAGAACATATGAACTGATCTTTTGACCTACCTTC 2820
QY 2819 AAGAAACCTTGGATCACGTAATGCTCTTAAAGTGAATAAATTTTGACCTGTGAGA 2878
Db 2821 AAGAAACCTTGGATCACGTAATGCTCTTAAAGTGAATAAATTTTGACCTGTGAGA 2880
QY 2879 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACA 2938
Db 2881 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACA 2940
QY 2939 GAAATGATCATACATTTTGATACCTGCGCATGTAAACATCTACTCTGAAATAAATGTGG 2998
Db 2941 GAAATGATCATACATTTTGATACCTGCGCATGTAAACATCTACTCTGAAATAAATGTGG 3000
QY 2999 TGCCATCAGGGGTACGGTTTGTGTAGTAACTTAATACCTTAACCCACATGCTCAA 3058
Db 3001 TGCCATCAGGGGTACGGTTTGTGTAGTAACTTAATACCTTAACCCACATGCTCAA 3060
QY 3059 AATCAATGATACATATCTCTGAGAGCCAGCAATACCAATGAAATTAATCAAAAAAAA 3118
Db 3061 AATCAATGATACATATCTCTGAGAGCCAGCAATACCAATGAAATTAATCAAAAAAAA 3120
QY 3119 AA 3120
Db 3121 AA 3122

RESULT 6
AX608735 LOCUS AX608735 4829 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 12 from Patent WO0231134.
ACCESSION AX608735
VERSION AX608735.1 GI:28404303
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
AUTHORS Qi,S., Akimanya,K.O., Riviere,P.J. and Junien,J.L.
TITLE Novel serine protease genes related to dppiv
JOURNAL Patent: WO 0231134-A 12 18-APR-2002;
Ferring BV (NL)
FEATURES
    source          Location/Qualifiers
    1..4829
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"

ORIGIN
Query Match          92.6%; Score 2889; DB 6; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 AAGTGTAAAGCTTCGAGGCCAAGCGCTGCTACTGCGCGCGCTGCTCTTTAGTGGCG 60
Db 1 AAGTGTAAAGCTTCGAGGCCAAGCGCTGCTACTGCGCGCGCTGCTCTTTAGTGGCG 60
QY 61 CGTTCGCGCGCTGGTGTGACGGCGCGCGCGAGGAAGCCACTGCAACAGGACCG 120
Db 61 CGTTCGCGCGCTGGTGTGACGGCGCGCGCGAGGAAGCCACTGCAACAGGACCG 120
```

```
QY 121 GAGTGGAGGGCGCGCGCAGCATGAAGCCGCGCAGGCCCGCTCCATAGCGCACGTCGCGGACGG 180
Db 121 GAGTGGAGGGCGCGCGCAGCATGAAGCCGCGCAGGCCCGCTCCATAGCGCACGTCGCGGACGG 180
QY 181 TCCGGGCGGGCGCGCGGGGGGAAAGAAAATGCAACATGGCAGCAGCAATGGAACAGAAACAG 240
Db 181 TCCGGGCGGGCGCGCGGGGGGAAAGAAAATGCAACATGGCAGCAGCAATGGAACAGAAACAG 240
QY 241 CTGGGTGTTGAGATATTTTGAACCTGCGGACTCTGAGGAGAAATPATTTGAATCAACGATCGG 300
Db 241 CTGGGTGTTGAGATATTTTGAACCTGCGGACTCTGAGGAGAAATPATTTGAATCAACGATCGG 300
QY 301 CTTAAATTTGAGGCGCTTTTATGTTGAGCGGTATTTCTCTGAGGTGAGCTTAAAGAGCTGCTT 360
Db 301 CTTAAATTTGAGGCGCTTTTATGTTGAGCGGTATTTCTCTGAGGTGAGCTTAAAGAGCTGCTT 360
QY 361 GCCGATACAGAAAAATATCATGGCTACATGATGGCTTAAGGGCACCATGATTTTATGTTTT 420
Db 361 GCCGATACAGAAAAATATCATGGCTACATGATGGCTTAAGGGCACCATGATTTTATGTTTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCCATGTCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTCTATTACCTTGGCCATGTCT 480
QY 481 GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540
QY 541 GCAGTCTTAATGCTCTCTTGGAGAGCCTCTTTTGGATCTTTTTCAGGGAACACTGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGAGCCTCTTTTGGATCTTTTTCAGGGAACACTGACTAT 600
QY 601 GGAATGATTTCTCGAGAGAGAGAACTATTAAGAGAGAAAGAAACGATTTGGAACAGTCGGA 660
Db 601 GGAATGATTTCTCGAGAGAGAGAACTATTAAGAGAGAGAAAGAAACGATTTGGAACAGTCGGA 660
QY 661 ATTGCTTCTTACGATTATCAACAGAGAGTGAAACATTTCTGTTTCAAGCGGTAGTGGGA 720
Db 661 ATTGCTTCTTACGATTATCAACAGAGAGTGAAACATTTCTGTTTCAAGCGGTAGTGGGA 720
QY 721 ATTTATCACGTAAAGATGAGGGGCCAAGAGATTACGCAACAACTTTTAAGGCCCAAT 780
Db 721 ATTTATCACGTAAAGATGAGGGGCCAAGAGATTACGCAACAACTTTTAAGGCCCAAT 780
QY 781 CTAGTGGAAACTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAAACTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA 840
QY 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACCCAGAGAA 900
Db 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACCCAGAGAA 900
QY 901 GAAAGGAGACTCACTTTATGCAATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTTATGCAATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960
QY 961 GCTGGAGTCTACTCTTTCTTCCAAAGAAATTTGATAGATATTTCTGGCTATTGTTGG 1020
Db 961 GCTGGAGTCTACTCTTTCTTCCAAAGAAATTTGATAGATATTTCTGGCTATTGTTGG 1020
QY 1021 TGTCCTCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATCTATATGAAGAA 1080
Db 1021 TGTCCTCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATCTATATGAAGAA 1080
QY 1081 AATGATGATCTGAGTGGAAATTTATCTATCATCCCTATGTTGGAAACAGAGAGG 1140
Db 1081 AATGATGATCTGAGTGGAAATTTATCTATCATCCCTATGTTGGAAACAGAGAGG 1140
QY 1141 GCAGATTCTTCGGTTATCTCTAAACAGAGTACAGAAATCTCTAAAGTCATCTTTTAAGATG 1200
Db 1141 GCAGATTCTTCGGTTATCTCTAAACAGAGTACAGAAATCTCTAAAGTCATCTTTTAAGATG 1200
```

QY 1201 TCAGAAATAATGATTGATGCTGAAGGAGGATCATAGATGTCTATAGATAAGGAACATAATT 1260  
DB 1201 TCAGAAATAATGATTGATGCTGAAGGAGGATCATAGATGTCTATAGATAAGGAACATAATT 1260  
QY 1261 CAACCTTTTGAGATTCTATTGTAAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCCT 1320  
DB 1261 CAACCTTTTGAGATTCTATTGTAAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCCT 1320  
QY 1321 GAGGGAATAATGCTTGGTCCATCTCTACTAGATGCTCCAGACTCGCCCTACAGATAGTG 1380  
DB 1321 GAGGGAATAATGCTTGGTCCATCTCTACTAGATGCTCCAGACTCGCCCTACAGATAGTG 1380  
QY 1381 TTGATCTCACTGAAATATTTATCCAGTAGAGATGATGTTATGGAAGGAGGAGAGCTC 1440  
DB 1381 TTGATCTCACTGAAATATTTATCCAGTAGAGATGATGTTATGGAAGGAGGAGAGCTC 1440  
QY 1441 ATTGAGTCAGTGCCTGATTCTGTGACGCCACTAATTAATCTATGAAGAAACAACAGACATC 1500  
DB 1441 ATTGAGTCAGTGCCTGATTCTGTGACGCCACTAATTAATCTATGAAGAAACAACAGACATC 1500  
QY 1501 TCGATAAATATCCATGACATCTTTTATGTTTTTCCCAAGTCAAGAGAGAAATGAG 1560  
DB 1501 TCGATAAATATCCATGACATCTTTTATGTTTTTCCCAAGTCAAGAGAGAAATGAG 1560  
QY 1561 TTTATTTTGGCTCTGAATGCAAAACAGTTTCCGTCATTTATACAAAATTAACATCTATT 1620  
DB 1561 TTTATTTTGGCTCTGAATGCAAAACAGTTTCCGTCATTTATACAAAATTAACATCTATT 1620  
QY 1621 TTAAGGAAGCAATATAACAGATCCAGTGTGGGTGCTGCTCCAGTAGATTTCAG 1680  
DB 1621 TTAAGGAAGCAATATAACAGATCCAGTGTGGGTGCTGCTCCAGTAGATTTCAG 1680  
QY 1681 TGTCTCTCAAGAGGAGATAGCAATTTACAGTGTGTAATGGGAGTTCTTCGCGCGGAT 1740  
DB 1681 TGTCTCTCAAGAGGAGATAGCAATTTACAGTGTGTAATGGGAGTTCTTCGCGCGGAT 1740  
QY 1741 GATCTCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800  
DB 1741 GATCTCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800  
QY 1801 TCCCTCTTTAGAGCATCACTGTAGTACAGTTAAGTAACTCTGAGAGGTGACAAGG 1860  
DB 1801 TCCCTCTTTAGAGCATCACTGTAGTACAGTTAAGTAACTCTGAGAGGTGACAAGG 1860  
QY 1861 CTGACTGACCGTGGCTACTACATTTCTTGCTGATCAGTCAGACCTGTGACTTTTATA 1920  
DB 1861 CTGACTGACCGTGGCTACTACATTTCTTGCTGATCAGTCAGACCTGTGACTTTTATA 1920  
QY 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCT 1980  
DB 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCT 1980  
QY 1981 GAAGATGCCCACTTGCAGAAACAAAGGAATTTTGGCCACCAATTTGGATTTCAGCAGGT 2040  
DB 1981 GAAGATGCCCACTTGCAGAAACAAAGGAATTTTGGCCACCAATTTGGATTTCAGCAGGT 2040  
QY 2041 CCTCTTCTGACTACTACTCTCCAGAAAATTTTCTCTTTGAAAGTACTACTGGATTACA 2100  
DB 2041 CCTCTTCTGACTACTACTCTCCAGAAAATTTTCTCTTTGAAAGTACTACTGGATTACA 2100  
QY 2101 TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTG 2160  
DB 2101 TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTG 2160  
QY 2161 CTGTTTCAATATATGGTGG--TCCTCAGGTGAGTTGGTGAATTAATCGTTTAAAGGAGTCA 2218  
DB 2161 CTGTTTCAATATATGGTGGTCTCTCAGGTGAGTTGGTGAATTAATCGTTTAAAGGAGTCA 2220  
QY 2219 AGTATTTCCGCTTGAAATACCTAGCTCTCTAGGTATGTTGTTAGTATAGACAACA 2278  
DB 2221 AGTATTTCCGCTTGAAATACCTAGCTCTCTAGGTATGTTGTTAGTATAGACAACA 2280  
QY 2279 GGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAA 2338

DB 2281 GGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAA 2340  
QY 2339 TAGAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATG 2398  
DB 2341 TAGAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATG 2400  
QY 2399 ACTTAGATCTGTGCGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAT 2458  
DB 2401 ACTTAGATCTGTGCGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAT 2460  
QY 2459 TAAATGACAGGTGAGATATCTTCAAGGTGCTATTCTGGGGCCCCCAGTCACTCTGTGGA 2518  
DB 2461 TAAATGACAGGTGAGATATCTTCAAGGTGCTATTCTGGGGCCCCCAGTCACTCTGTGGA 2520  
QY 2519 TCTTCTATGATACAGGATACACGGAACTGTTATATGGGTCACTGCTGACGAAATGAACAGG 2578  
DB 2521 TCTTCTATGATACAGGATACACGGAACTGTTATATGGGTCACTGCTGACGAAATGAACAGG 2580  
QY 2579 GCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTT 2638  
DB 2581 GCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTT 2640  
QY 2639 TACTGCTCTTACATGTTTCTTGGATGAGAAATGTCATTTTGACATACCAAGTATATTAC 2698  
DB 2641 TACTGCTCTTACATGTTTCTTGGATGAGAAATGTCATTTTGACATACCAAGTATATTAC 2700  
QY 2699 TGAGTTTTTTTATGTCAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGAC 2758  
DB 2701 TGAGTTTTTTTATGTCAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGAC 2760  
QY 2759 ACAGCATAAAGAGTTTCTCGAATCGGGAGAAACATTTATGAACTGATCTTTTGCACTACCTTC 2818  
DB 2761 ACAGCATAAAGAGTTTCTCGAATCGGGAGAAACATTTATGAACTGATCTTTTGCACTACCTTC 2820  
QY 2819 AAGAAACCTTTGGATCAGTATTCTGCTCTAAAAGTATATAATTTTGACCTGTGTAGA 2878  
DB 2821 AAGAAACCTTTGGATCAGTATTCTGCTCTAAAAGTATATAATTTTGACCTGTGTAGA 2880  
QY 2879 ACTCTCTGATATACCTGGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAACACA 2938  
DB 2881 ACTCTCTGATATACCTGGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAACACA 2940  
QY 2939 GAAATGATCATCACTTTTGTATACCTGCGATGAAACATCTACTCTGGAATAAATGTGG 2998  
DB 2941 GAAATGATCATCACTTTTGTATACCTGCGATGAAACATCTACTCTGGAATAAATGTGG 3000  
QY 2999 TGCCATGCAAGGGCTACGGTTTGTGTATCTAATACCTTAAACCCCATCTCAA 3058  
DB 3001 TGCCATGCAAGGGCTACGGTTTGTGTATCTAATACCTTAAACCCCATCTCAA 3060  
QY 3059 AATCAATGATACATATTTCTTGAGAGACCCAGCAATACCATAAGAAATTTACTTAAAAAAA 3118  
DB 3061 AATCAATGATACATATTTCTTGAGAGACCCAGCAATACCATAAGAAATTTACTTAAAAAAA 3120  
QY 3119 AA 3120  
DB 3121 AA 3122

RESULT 7  
BC040203

DEFINITION Homo sapiens dipeptidylpeptidase 8, mRNA (cDNA clone MGC:42570 IMAGE:4824813), complete cds.  
ACCESSION BC040203  
VERSION BC040203.1 GI:26007916  
KEYWORDS MGC:.,.,.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.



REFERENCE	1 (bases 1 to 4535)	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F., Hopkins,R.F., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Diatchenko,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Stapleton,L., Brownstein,M.J., Uedin,T.B., Toshituki,S., Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Toshituki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Sanchez,J., Helton,E., Kettelman,M., Madan,A.C., Rodrigues,S., Fahey,J., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakeley,R.W., Trumbleman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Groudman,J.J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E., Schnarch,A., Schein,J.E., Jones,S.J., and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL	Proc Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
PUBMED	12477932	
REFERENCE	2 (bases 1 to 4535)	
AUTHORS	Strausberg,R.	
TITLE	Direct Submission	
JOURNAL	Submitted (27-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgabs-r@mail.nih.gov">cgabs-r@mail.nih.gov</a> Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a> Contact: <a href="mailto:amg@bcm.tmc.edu">amg@bcm.tmc.edu</a> Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.	
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAP Plate 71 Row: k Column: 23 Location/Qualifiers 1..4535 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:42570 IMAGE:4824813" /tissue_type="Testis" /clone_lib="NIH MGC 97" /lab_host="DH10B" /notes="Vector: pBluescript" 1..4535 /gene="DPP8" /notes="synonyms: FLJ20283, FLJ14920, MGC26191" /db_xref="GeneID:54878" /db_xref="MIM:606819" 1629..4277 /gene="DPP8" /codon_start=1 /product="Dpp8 protein" /protein_id="AAH40203.1" /db_xref="GI:26007917" /db_xref="GeneID:54878" /db_xref="MIM:606819" /translation="MAAAMETSLGVEIFETADCEENIESODRPKLPPFVVERYSWS	
gene		
CDS		

Db 2276 TGGCAACGATATTTTGGATATCTAACATCGTAAACGAGAGAAAGGAGACTCACTTATGT 2335  
Qy 921 GCACATCAGCTAGCCACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 980  
Db 2336 GCACATCAGCTAGCCACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 2395  
Qy 981 TCTCCAGAAGAAATTTGATAGATATCTGGCTATTGGTGGTGTCCAAAGAGCTGAAACAAC 1040  
Db 2396 TCTCCAGAAGAAATTTGATAGATATCTGGCTATTGGTGGTGTCCAAAGAGCTGAAACAAC 2455  
Qy 1041 TCCAGTGGTGGTAAAAATCTTAGAATTTCTATATGAAGAAAAATGATGAATCTGAGGTGGA 1100  
Db 2456 TCCAGTGGTGGTAAAAATCTTAGAATTTCTATATGAAGAAAAATGATGAATCTGAGGTGGA 2515  
Qy 1101 AATTATTCATGTTACATCCCTATGTTGGAAAAACAAGAGGGCAGATTCATTTCCGTTATCC 1160  
Db 2516 AATTATTCATGTTACATCCCTATGTTGGAAAAACAAGAGGGCAGATTCATTTCCGTTATCC 2575  
Qy 1161 TAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAGATGTCAGAAAAATGATGATGC 1220  
Db 2576 TAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAGATGTCAGAAAAATGATGATGC 2635  
Qy 1221 TGAAGAGGATCATAGATGTCATAGATAGGAATTAATTTCAACCTTTTGAGATTCATTT 1280  
Db 2636 TGAAGAGGATCATAGATGTCATAGATAGGAATTAATTTCAACCTTTTGAGATTCATTT 2695  
Qy 1281 TGAAGAGGATTCGAATATATGTCAGAGCTGGATGGAGTCTCTGAGGGAAATATGCTTGTC 1340  
Db 2696 TGAAGAGGATTCGAATATATGTCAGAGCTGGATGGAGTCTCTGAGGGAAATATGCTTGTC 2755  
Qy 1341 CATCTCTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTTGGAATTAAT 1400  
Db 2756 CATCTCTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTTGGAATTAAT 2815  
Qy 1401 TATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATTTAGTCAAGTGCCTGATTC 1460  
Db 2816 TATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATTTAGTCAAGTGCCTGATTC 2875  
Qy 1461 TGTGACGCCATTAATTTATATGAGAAACAACAGACATCTCGATTAATATCCATGACAT 1520  
Db 2876 TGTGACGCCATTAATTTATATGAGAAACAACAGACATCTCGATTAATATCCATGACAT 2935  
Qy 1521 CTTTCATGTTTTTCCCCAAGTCACGAGAGGAAATTTGATTTATTTTGGCTCTGATG 1580  
Db 2936 CTTTCATGTTTTTCCCCAAGTCACGAGAGGAAATTTGATTTATTTTGGCTCTGATG 2995  
Qy 1581 CAAAAACAGTTTTCCGTCAATTTATACAAAAATTAATCATCTATTTTAAAGGAAAGCAATATAA 1640  
Db 2996 CAAAAACAGTTTTCCGTCAATTTATACAAAAATTAATCATCTATTTTAAAGGAAAGCAATATAA 3055  
Qy 1641 AGATTCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAAAGAGGAGAT 1700  
Db 3056 AGATTCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAAAGAGGAGAT 3115  
Qy 1701 AGCAATTAACAGTGGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCCAAAGTTGA 1760  
Db 3116 AGCAATTAACAGTGGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCCAAAGTTGA 3175  
Qy 1761 TGAAGTCAGAAGCTGGTGTATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCT 1820  
Db 3176 TGAAGTCAGAAGCTGGTGTATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCT 3235  
Qy 1821 GTACGTAGTCAAGTGAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTC 1880  
Db 3236 GTACGTAGTCAAGTGAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTC 3295  
Qy 1881 ACATTCCTGCTGCATCAGTCAGCACTGTGACTTTCTTTTAAAGTATAGTATACCCAGAA 1940  
Db 3296 ACATTCCTGCTGCATCAGTCAGCACTGTGACTTTCTTTTAAAGTATAGTATAGTATACCCAGAA 3355  
Qy 1941 GAATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 2000  
Db 3356 GAATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA 3415

Qy 2001 AACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTCTGACTATACTCC 2060  
Db 3416 AACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTCTGACTATACTCC 3475  
Qy 2061 TCCAGAAATTTTCTCTTTTGAAGTACTACTGCGATTTTCAATTTGATGGGATCTCTACAA 2120  
Db 3476 TCCAGAAATTTTCTCTTTTGAAGTACTACTGCGATTTTCAATTTGATGGGATCTCTACAA 3535  
Qy 2121 GCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGCTGTTCTCATATATGGTGGTCC 2180  
Db 3536 GCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGCTGTTCTCATATATGGTGGTCC 3595  
Qy 2181 TCAGTGCAGTTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 2240  
Db 3596 TCAGTGCAGTTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 3655  
Qy 2241 AGCCTCTCTAGGTTATGCTGGTGTAGTAGACACAGGGGATCTCTGTCACCGAGGCT 2300  
Db 3656 AGCCTCTCTAGGTTATGCTGGTGTAGTAGACACAGGGGATCTCTGTCACCGAGGCT 3715  
Qy 2301 TAAATTTGAAGCGCCTTTAAATATATAAATGGGTCAAAATAGAAAATTCAGCATCAGGTGA 2360  
Db 3716 TAAATTTGAAGCGCCTTTAAATATATAAATGGGTCAAAATAGAAAATTCAGCATCAGGTGA 3775  
Qy 2361 AGGACTCCAATPATCTAGCTTCTCGATATGATTTCAATTTGATCTGTTAGATCTGTTGGGATCCA 2420  
Db 3776 AGGACTCCAATPATCTAGCTTCTCGATATGATTTCAATTTGATCTGTTAGATCTGTTGGGATCCA 3835  
Qy 2421 CGGCTGGTCTTATGAGGAGTACCTCTCCCTGATGGCATTTAAATGCAGAGGTCAAGATCTT 2480  
Db 3836 CGGCTGGTCTTATGAGGAGTACCTCTCCCTGATGGCATTTAAATGCAGAGGTCAAGATCTT 3895  
Qy 2481 CAGGGTGTCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 2540  
Db 3896 CAGGGTGTCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 3955  
Qy 2541 GGAACGTATATATGGTTCACCTGACAGAGTGAACAGGGCTATTTAGGATCTGTGGC 2600  
Db 3956 GGAACGTATATATGGTTCACCTGACAGAGTGAACAGGGCTATTTAGGATCTGTGGC 4015  
Qy 2601 CATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTCTCT 2660  
Db 4016 CATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTCTCT 4075  
Qy 2661 GGATGAGAAATGTCATTTTGCACATACCAAGTATTTACTGAGTTTTTTAGTGAAGGCTGG 2720  
Db 4076 GGATGAGAAATGTCATTTTGCACATACCAAGTATTTACTGAGTTTTTTAGTGAAGGCTGG 4135  
Qy 2721 AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGAGTTCTTGAATC 2780  
Db 4136 AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGAGTTCTTGAATC 4195  
Qy 2781 GGGAGAACATTTATGAACTGCACTTTTGGACTACTCTTCAAGAAAACCTTGGATCAGGTAT 2840  
Db 4196 GGGAGAACATTTATGAACTGCACTTTTGGACTACTCTTCAAGAAAACCTTGGATCAGGTAT 4255  
Qy 2841 TGCTGCTCTAAAGTGATATAAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTA 2900  
Db 4256 TGCTGCTCTAAAGTGATATAAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTA 4315  
Qy 2901 TTTTAAACCAATGAGGAGGTTTTTAATCAACAGAAAACAGAAATTTGATCATCATTTTGTAT 2960  
Db 4316 TTTTAAACCAATGAGGAGGTTTTTAATCAACAGAAAACAGAAATTTGATCATCATTTTGTAT 4375  
Qy 2961 ACCTGCCATGTAACATCTACTCTCTGAAAATTAATTTGGTGCCTATGAGGGGTCTACGGTT 3020  
Db 4376 ACCTGCCATGTAACATCTACTCTCTGAAAATTAATTTGGTGCCTATGAGGGGTCTACGGTT 4435  
Qy 3021 TGTGGTAGTAAATCTTAATACCTTTAACCCCAATGCTCAAAATCAAAATGATATATTTCTG 3080  
Db 4436 TGTGGTAGTAAATCTTAATACCTTTAACCCCAATGCTCAAAATCAAAATGATATATTTCTG 4495



QY	3081	AGAGACCCAGCAATACCATAGAAATTACTATAAAAAAAAAAAAAA	3120
Db	4496	AGAGACCCAGCAATACCATAGAAATTACTATAAAAAAAAAAAAAA	4535
RESULT 8			
LOCUS	AY354202	3030 bp mRNA linear	PRI 02-SEP-2003
DEFINITION	Homo sapiens dipeptidyl peptidase 8 isoform 3 (DPP8) mRNA, complete cds, alternatively spliced.		
ACCESSION	AY354202		
VERSION	AY354202.1	GI:34329351	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.		
REFERENCE	1 (bases 1 to 3030)		
AUTHORS	Sha, J.H., Zhou, Z.M. and Li, J.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-JUL-2003) Key Lab of Reproductive Medicine, Nanjing Medical University, 140 Han Zhong Road, Nanjing, Jiangsu 210029, China		
FEATURES			
source	Location/Qualifiers		
	1..3030		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="15"		
	/tissue_type="testis"		
gene	1..3030		
	/gene="DPP8"		
CDS	161..2857		
	/gene="DPP8"		
	/note="alternatively spliced"		
	/codon_start=1		
	/product="dipeptidyl peptidase 8 isoform 3"		
	/protein_id="AAQ63887.1"		
	/db_xref="GI:34329352"		
	/translation="MKRSEOMKIKSGKNMAAMETEQLGVIEFTADCEENIESQD RPKEFPYVERYSQKLLADTRKYHYGMKAPHDFMFKRNDPDGPHSDRIYYL AMSEGNEDFYSEIPKTIINRAAVLMISWPLDLFOALDIDYGMREBELLRERK IGTVGLASVDYHSGSTFLFQAGSGIYHVKDGGQGFQQLRPNLVETSCPNRMIDP KLCPPADPDWIAFTHSNDIWSINVTREERLLTYVHNELANNEEDARAGVATFLQBE FDRYSGVWCPCKAETTPSGKILRLIYEENDESEVEIITHVSPMLETRRADSPRYEKE GTANPKVTFKMSIMIDAEGRIDVIDKELIOPPEILFEGVEIYARAGMTPEKVAWS ILDRSOTRLOIIVLISPELPIVEDDVMERORLIESVPSVTPLIIVEETDIDWNIH DIHPVQSHREBIEFTFASECKTGRHLVYKITSILKSKYKRSGGLPAPSPKCIPI KEELATISGMEVLEIRHGSNIQVDSRRLVYFEGTKDSPLSHLYVYVYNPGEVTRL TDRYSHSCCISOHCDFFISKYSNKNPHCVSLYKLSPEDDPTCKTKEFWATILDSA GPLPDTPPIEFSEFTGTGLYMLYKPHDLQPKGYPTVLFYGGPQVQLVNNRFK GVXYFLRLTLASGYVYVVDNRGSRHLKFEAGFKYKMSQIEIDDPQVGLQYLAR YDFIDLDRLVGHGWSYGGVLSLMALMQBSDLFRVALAGAPVTLWIFDYTCYTRYNGH PDONEQGYLIGSVAMQAEKFPSEPNRLILLHGFIDENVHFHTSILLSFLVRAGKPYD LQIYPOERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI"		
ORIGIN			
Query Match 88.3%; Score 2756; DB 8; Length 3030;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2806; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	201	AGGAAATGCAACATGCGACGACGCAATGGAACAGAACGCTGGGTGTGAGATATTGA	260
Db	196	AGGAAATGCAACATGCGACGACGCAATGGAACAGAACGCTGGGTGTGAGATATTGA	255
QY	261	AACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGAGCCTTTTA	320
Db	256	AACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGAGCCTTTTA	315
QY	321	TGTTGAGCGGTATTCTCGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAAAATCA	380
Db	316	TGTTGAGCGGTATTCTCGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAAAATCA	375

QY 1461 TGTGACGCCAATAATATCTATGAGAAACACAGACATCTGGATAAATATCCATGACAT 1520  
DB 1456 TGTGACGCCAATAATATCTATGAGAAACACAGACATCTGGATAAATATCCATGACAT 1515  
QY 1521 CTTTCATGTTTTCGCCCAAGTCCACGAGAGGAATTCAGTTATTTTTCGCTCTGAATG 1580  
DB 1516 CTTTCATGTTTTCGCCCAAGTCCACGAGAGGAATTCAGTTATTTTTCGCTCTGAATG 1575  
QY 1581 CAAACACAGTTTCGCTCATTTATACAAATTTACATCTATTTTAAAGGAAACCAATATAA 1640  
DB 1576 CAAACACAGTTTCGCTCATTTATACAAATTTACATCTATTTTAAAGGAAACCAATATAA 1635  
QY 1641 ACGATCCAGTGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAAAGAGGAGAT 1700  
DB 1636 ACGATCCAGTGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAAAGAGGAGAT 1695  
QY 1701 AGCAATTTACAGTGTGGTGAATGGGAAGTTCTTGCGCGGATGATCTAATATCCAAAGTTGA 1760  
DB 1696 AGCAATTTACAGTGTGGTGAATGGGAAGTTCTTGCGCGGATGATCTAATATCCAAAGTTGA 1755  
QY 1761 TGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCT 1820  
DB 1756 TGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCT 1815  
QY 1821 GTACGTAGTCAGTTACGTAAATCCTCGAGAGGTGACAGGCTGACTGACCGTGCTACTC 1880  
DB 1816 GTACGTAGTCAGTTACGTAAATCCTCGAGAGGTGACAGGCTGACTGACCGTGCTACTC 1875  
QY 1881 ACATTTCTGTCATCAGTCAGCAGTCGACTTCTTTTAAAGTAAAGTATAGTACACAGAA 1940  
DB 1876 ACATTTCTGTCATCAGTCAGCAGTCGACTTCTTTTAAAGTAAAGTATAGTACACAGAA 1935  
QY 1941 GAATCCACACTGTGTGCTCCCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCAACTGCAA 2000  
DB 1936 GAATCCACACTGTGTGCTCCCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCAACTGCAA 1995  
QY 2001 AACAAAGGAATTTTGGGCCACCAATTTTGAATTCAGCAGGTCCTCTTCCTGACTATATCC 2060  
DB 1996 AACAAAGGAATTTTGGGCCACCAATTTTGAATTCAGCAGGTCCTCTTCCTGACTATATCC 2055  
QY 2061 TCAGAAATTTCTCTTTTGAAGTACTA CTGGATTTACATGTATGGATGCTCTACAA 2120  
DB 2056 TCAGAAATTTCTCTTTTGAAGTACTA CTGGATTTACATGTATGGATGCTCTACAA 2115  
QY 2121 GCCTCATGATCTACAGCTGGAAAGAAATATCTCTACTGCTGTGTTTCATATATGCTGCTC 2180  
DB 2116 GCCTCATGATCTACAGCTGGAAAGAAATATCTCTACTGCTGTGTTTCATATATGCTGCTC 2175  
QY 2181 TCAGGTCAGTGTGGTGAATTCGTTTAAAGAGTCAAGTATTTTCGCTTGAATACCCCT 2240  
DB 2176 TCAGGTCAGTGTGGTGAATTCGTTTAAAGAGTCAAGTATTTTCGCTTGAATACCCCT 2235  
QY 2241 AGCTCTCTAGGTTATGTGGTGTAGTATAGACACAGGGGATCCTGTCAACCGAGGCT 2300  
DB 2236 AGCTCTCTAGGTTATGTGGTGTAGTATAGACACAGGGGATCCTGTCAACCGAGGCT 2295  
QY 2301 TAAATTTGAAGCGCCCTTTAAATATATAAATGGGTCAAAATAGAAATTCAGTACAGGTGGA 2360  
DB 2296 TAAATTTGAAGCGCCCTTTAAATATATAAATGGGTCAAAATAGAAATTCAGTACAGGTGGA 2355  
QY 2361 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTCATGATCGTGTGGGATCCCA 2420  
DB 2356 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTCATGATCGTGTGGGATCCCA 2415  
QY 2421 CAGCTGTGCTTATGGAGGATACCTCTCCCTGATGGCATTTAATGACAGGTCAGATATCTT 2480  
DB 2416 CAGCTGTGCTTATGGAGGATACCTCTCCCTGATGGCATTTAATGACAGGTCAGATATCTT 2475  
QY 2481 CAGGGTGTCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACAC 2540  
DB 2476 CAGGGTGTCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACAC 2535  
QY 2541 GGAACGTTATATGGGTCACCTCGACCAAGATGAACAGGCTATTACTTTAGGATCTGTGGC 2600

DB 2536 GGAACGTTATATGGGTCACCTTGACCAAGTGAACAGGCTATTACTTAGGATCTGTGGC 2595  
QY 2601 CATGCAAGCAGAAAAGTTCCCTCTGAAACCAATGTTTACTGCTCTTACATGTTTCTCT 2660  
DB 2596 CATGCAAGCAGAAAAGTTCCCTCTGAAACCAATGTTTACTGCTCTTACATGTTTCTCT 2655  
QY 2661 GGAATGAGAATGTCATTTTGCACATACCATGATATTTACTGAGTTTTTTAGTGAGGCTGG 2720  
DB 2656 GGAATGAGAATGTCATTTTGCACATACCATGATATTTACTGAGTTTTTTAGTGAGGCTGG 2715  
QY 2721 AAAACCATATGATTTACAGATCTATCTCAGGAGAGACACAGATTAAGATTCCTGAATC 2780  
DB 2716 AAAACCATATGATTTTACAGATCTATCTCAGGAGAGACACAGATTAAGATTCCTGAATC 2775  
QY 2781 GGGAGAACATTAAGAACTGATCTTTTGCACACTCTTCAAGAAACCTTGGATCAAGTAT 2840  
DB 2776 GGGAGAACATTAAGAACTGATCTTTTGCACACTCTTCAAGAAACCTTGGATCAAGTAT 2835  
QY 2841 TGCTGCTCTAAAAGTGAATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTA 2900  
DB 2836 TGCTGCTCTAAAAGTGAATATAATTTTGAACCTGTGTAGAACTCTCTGGTATACACTGGCTA 2895  
QY 2901 TTTTAAACCAATAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATTTTGGAT 2960  
DB 2896 TTTTAAACCAATAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATTTTGGAT 2955  
QY 2961 ACCTGCCATGTAACATCTACTCTCGAAATAAATGTTGGTGCATGCA 3007  
DB 2956 ACCTGCCATGTAACATCTACTCTCGAAATAAATGTTGGTGCATGCA 3002

RESULT 9  
AX354795  
LOCUS AX354795 2649 bp DNA linear PAT 06-FEB-2002  
DEFINITION Sequence 3 from Patent WO0179473.  
ACCESSION AX354795  
VERSION AX354795.1 GI:18619528  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
1  
Meyers, R.A. and Williamson, M.  
21953, a human prolyl oligopeptidase family member and uses thereof  
Patent: WO 0179473-A 3 25-Oct-2001;  
Millennium Pharmaceuticals, Inc. (US)

FEATURES  
source  
1..2649  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 81.6%; Score 2547; DB 6; Length 2649;  
Best Local Similarity 99.9%; Pred. No; 0;  
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTTAACTGCGGACTGT 273  
DB 1 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTTAACTGCGGACTGT 60  
QY 274 GAGGAGATATTGAATCAGAGGATCGGCTTAAATTTGAGGCTTTTATGTTGAGCGGTAT 333  
DB 61 GAGGAGATATTGAATCAGAGGATCGGCTTAAATTTGAGGCTTTTATGTTGAGCGGTAT 120  
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACCAAGAAATATCATGGCTTACATGATG 393  
DB 121 TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACCAAGAAATATCATGGCTTACATGATG 180  
QY 394 GCTAAGGCCACCATGATTTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATTC 453

Db GCTAAGGCCACCAATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 240  
Qy GACAGATCTATTACCTTGCCATGCTCGTGTGAGAACAGAGAAATACACTGTTTATCT 513  
Db GACAGATCTATTACCTTGCCATGCTCGTGTGAGAACAGAGAAATACACTGTTTATCT 300  
Qy GAAATTCCTCAAACTATCAATAGAGCAGCAGCTTAAATGCTCTCTTGGAAGCCCTCTTTG 573  
Db GAAATTCCTCAAACTATCAATAGAGCAGCAGCTTAAATGCTCTCTCTTGGAAGCCCTCTTTG 360  
Qy GATCTTTTCAGGCAACATCGGACTATGGAATGTATTCGAGAGAAAGAACTATTAGA 633  
Db GATCTTTTCAGGCAACATCGGACTATGGAATGTATTCGAGAGAAAGAACTATTAGA 420  
Qy GAAAGAAAACGCAATGGAAACAGTCGGAATGTCTTACGATTAATCAACAGGAAGTGA 693  
Db GAAAGAAAACGCAATGGAAACAGTCGGAATGTCTTACGATTAATCAACAGGAAGTGA 480  
Qy ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTAAGATGGAAGGCCCAAGGA 753  
Db ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTAAGATGGAAGGCCCAAGGA 540  
Qy TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGGATG 813  
Db TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCAACATACGGATG 600  
Qy GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATTT 873  
Db GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATTT 660  
Qy TGSATNTAATCATCGTAACAGAGAAAGAGAGACTACATATATGTCACATAGACTA 933  
Db TGSATNTAATCATCGTAACAGAGAAAGAGAGACTACATATATGTCACATAGACTA 720  
Qy GCCAACHATGAAGAGATGCGAGTCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAAGAA 993  
Db GCCAACHATGAAGAGATGCGAGTCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAAGAA 780  
Qy TTTGATAGATATTCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCACAGTGGTGGT 1053  
Db TTTGATAGATATTCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCACAGTGGTGGT 840  
Qy AAAATCTTTAGAAATCTTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113  
Db AAAATCTTTAGAAATCTTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 900  
Qy ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 1173  
Db ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 960  
Qy GCAATCCTAAAGTCACCTTTAAGATGTCAGAAATAATGATTTGATGCTGAAGGAGATC 1233  
Db GCAATCCTAAAGTCACCTTTAAGATGTCAGAAATAATGATTTGATGCTGAAGGAGATC 1020  
Qy ATAGATGTCATAGTAAGNACTAATTCACACCTTTTGAGATTCATTTGAAGGATGGA 1293  
Db ATAGATGTCATAGTAAGNACTAATTCACACCTTTTGAGATTCATTTGAAGGATGGA 1080  
Qy TATATGCCAGAGCTGGATGGACTCTGAGGAAATATGCTGGTCCATCTACTAGAT 1353  
Db TATATGCCAGAGCTGGATGGACTCTGAGGAAATATGCTGGTCCATCTACTAGAT 1140  
Qy CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413  
Db CGCTCCAGACTCGCTCGAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1200  
Qy GATGATGTTATGAAAGGAGAGACTCATTTGAGTCAGTGCCTGATCTGTGAGCCACTA 1473  
Db GATGATGTTATGAAAGGAGAGACTCATTTGAGTCAGTGCCTGATCTGTGAGCCACTA 1260  
Qy ATTAATCTATGAAGAAAACACAGACATCTGGATGAATATCCATGACATCTTTTCATGTTTTT 1533

Db ATTATCTATGAAGAAAACACAGACATCTCGATAAATATCCATGACATCTTTCACTGTTTTT 1261  
Qy CCCAAAGTCA CGAAGAGAAATTTGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTTT 1593  
Db CCCAAAGTCA CGAAGAGAAATTTGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTTT 1380  
Qy CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAAAACGATCCAGTGGT 1653  
Db CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAAAACGATCCAGTGGT 1440  
Qy GGGCTGCTCTCCAAAGTGTCTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCACT 1713  
Db GGGCTGCTCTCCAAAGTGTCTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCACT 1500  
Qy GGTCAATGGGAAGTTCTTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAAG 1773  
Db GGTCAATGGGAAGTTCTTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAAG 1560  
Qy CTGGTATATTTTCAAGGCAACCAAGACTCCCTCTTTAGAGCATCACTCTGATAGTCACT 1833  
Db CTGGTATATTTTCAAGGCAACCAAGACTCCCTCTTTAGAGCATCACTCTGATAGTCACT 1620  
Qy TACGTAATCTCGAGAGGTGACAGGCTGACCTGACCTGGCTACTACATTTCTGCTGC 1893  
Db TACGTAATCTCGAGAGGTGACAGGCTGACCTGACCTGGCTACTACATTTCTGCTGC 1680  
Qy ATCAGTCAGCACTGACCTCTTATTAAGTAAGTATAGTAAACAGAGATCCACACTGT 1953  
Db ATCAGTCAGCACTGACCTCTTATTAAGTAAGTATAGTAAACAGAGATCCACACTGT 1740  
Qy GTGTCCCTTTACAAAGCTATCAAGTCTCTGAAAGTACCCAACTTGCAAAACAAAGAAATTT 2013  
Db GTGTCCCTTTACAAAGCTATCAAGTCTCTGAAAGTACCCAACTTGCAAAACAAAGAAATTT 1800  
Qy TGGGCCACCAATTTTGGATTCAGGAGTCTCTCTCTGACTATACCTTCCAGAAATTTTTC 2073  
Db TGGGCCACCAATTTTGGATTCAGGAGTCTCTCTCTGACTATACCTTCCAGAAATTTTTC 1860  
Qy TCTTTTGAAGTACTACTCGATTTACATTTGATGGAATGCTCTCAAGACCTCATGATCTA 2133  
Db TCTTTTGAAGTACTACTCGATTTACATTTGATGGAATGCTCTCAAGACCTCATGATCTA 1920  
Qy CAGCTCGGAAGAAATATCCTACTGTCTGTTTATATATGTTGGTGGTCTCAGGTGCAGTTG 2193  
Db CAGCTCGGAAGAAATATCCTACTGTCTGTTTATATATGTTGGTGGTCTCAGGTGCAGTTG 1980  
Qy GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTTGAATACCTAGCCTCTCTAGGT 2253  
Db GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTTGAATACCTAGCCTCTCTAGGT 2040  
Qy TATGTGTTGTAGTATAGCAACAGGGGATCCTGTACCGAGGCTTAAATTTGAAGGC 2313  
Db TATGTGTTGTAGTATAGCAACAGGGGATCCTGTACCGAGGCTTAAATTTGAAGGC 2100  
Qy GCCTTTAAATATAAATGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCAATAT 2373  
Db GCCTTTAAATATAAATGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCAATAT 2160  
Qy CTAGCTTCTCGATATGATTTCACTGACTAGTACCTGTGGGATCCAGGCTGGCTCTAT 2433  
Db CTAGCTTCTCGATATGATTTCACTGACTAGTACCTGTGGGATCCAGGCTGGCTCTAT 2220  
Qy GAGGATACCTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTCAGGGTGGCTATT 2493  
Db GAGGATACCTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTCAGGGTGGCTATT 2280  
Qy GCTGGGCCCCAGTCACTCTGTGGATCTTCTATATAGATACAGGATACAGGAACTGATATG 2553  
Db GCTGGGCCCCAGTCACTCTGTGGATCTTCTATATAGATACAGGATACAGGAACTGATATG 2340  
Qy GGTCACTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGGATCTGCAAGCAGAA 2613  
Db GGTCACTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGGATCTGCAAGCAGAA 2400

2614 AAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTC 2673  
|||||  
2401 AAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTC 2460  
|||||  
2674 CATTTTGACATACAGTATATTAAGTGGTTTATAGTGGGCTGGAAAGCCATATGAT 2733  
|||||  
2461 CATTTTGACATACAGTATATTAAGTGGTTTATAGTGGGCTGGAAAGCCATATGAT 2520  
|||||  
2734 TTACAGATCTATCTCAGGAGACACAGCATAGAGTTCTTGAATCGGAGACATATAT 2793  
|||||  
2521 TTACAGATCTATCTCAGGAGACACAGCATAGAGTTCTTGAATCGGAGACATATAT 2580  
|||||  
2794 GAATGCACTCTTTGCACTACCTTCAAGAAACCTTTGGATCAGTATGCTGCTCTAAAA 2853  
|||||  
2581 GAATGCACTCTTTGCACTACCTTCAAGAAACCTTTGGATCAGTATGCTGCTCTAAAA 2640  
|||||  
2854 GTGATATAA 2862  
|||||  
2641 GTGATATAA 2649

## RESULT 10

LOCUS AY172659 2656 bp mRNA linear PRI 08-JAN-2003  
DEFINITION Homo sapiens dipeptidyl peptidase IV-related protein-1 (DPRP1)  
mRNA, complete cds.

## ACCESSION

AY172659

## VERSION

AY172659.1 GI:27549549

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 2656)  
Qi,S., Akiananya,K., Riviere,P. and Junien,J.-L.  
Novel Serine Protease Genes Related To DPPIV  
Patent: US (WO 0231134)-A 18-APR-2002;  
2 (bases 1 to 2656)  
Qi,S., Akiananya,K., Riviere,P. and Junien,J.-L.  
Direct Submission  
Submitted (04-NOV-2002) Ferring Research Institute, 3550 General  
Atomics Ct., San Diego, CA 92121, USA

## FEATURES

## source

1..2656  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q22.1-q22.2"  
1..2656  
/gene="DPRP1"  
8..2656  
/gene="DPRP1"  
/codon\_start=1  
/product="dipeptidyl peptidase IV-related protein-1"  
/protein\_id="AA017261.1"  
/db\_xref="GI:27549550"  
/translation="MAAAMETEQLGVETPFDADCEINTESQDRPKLPPFFVVERYSWSQ  
LKLLADTRHYGMMAKAPHFVFRNDPDGHSRIYYLAWSGNRENTLYFSI  
PTFNRAALVLMKSWKLLDQFADLYGMYREELLRRKRI GTVGIASVDYHQSG  
TFIQAAGSYHYVKGDPQGTQOPLRPNLVTSCPNTRMDPKLCPADPDWIAFHNS  
DIWLSNVTREERLLTYVHNELANNEEDARSAGVATFVLQSEFDYSGYWCPEATT  
PSGKIRULIENDESEVEIIVHTVSPMLSTRADSFPRIKPTGANPKVTFKMSLMI  
DAEGRIDVDKELI QFFELFEGVEYIARAGWTPGKYAWSILLDRSQRLQVLIS  
PELPIVEDDWMQRQLIESVPSVPLI IYEETPD IWINIHD IEFVFPQSHBEIEF  
IFASEGTRFRLHYKITSILKSKYKRSGGLPAPSPDKPIKEIATITGEVEIAR  
HGSNIQVDEVRVLVEPSTKDSPLLEHLVYVSNVNGEVLTDRGYSHCSCI SOHCD  
FTTISKYQNQKPCVSLYKLSSPEDDPTCKEFWATILDSAGLPDYTPPEIFSFES  
VVVIDNRGSCHRGLKPEGAFKVMQGI EIDDQVEGLQYLAIRYDFIDLDVRVGHWSY  
GGYLSLMALMQRSDI FRVAIAGAPVTLWIFDYTGTRYMGHPQNEQGYVLSVAMQ

## gene

## CDS

AEKPPSPBNRLLLHLGHFLDENVHAFTSILLSFLVRAGKPYDLQIYPOERHSIRVPES  
GEHYELHLLHYLQENLGSRIAALKVI"

## ORIGIN

Query Match 80.0%; Score 2495; DB 8; Length 2656;  
Best Local Similarity 99.9%; Pred. No! 0;  
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
213 CATGCGACGCAATGGAACAGAACAGCTGGGTGTGAGATATTTGAAACTCGCGACTG 272  
Db 7 CATGCGACGCAATGGAACAGAACAGCTGGGTGTGAGATATTTGAAACTCGCGACTG 66  
273 TGAGGAGATATTTGAATCAGAGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTA 332  
Db 67 TGAGGAGATATTTGAATCAGAGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTA 126  
333 TTCCTGGAGTCAGCTTAAAAAGCTTGGCCGATACCAGAAATATCATGGCTACATGAT 392  
Db 127 TTCCTGGAGTCAGCTTAAAAAGCTTGGCCGATACCAGAAATATCATGGCTACATGAT 186  
393 GGCTAAGGCACCATGATTTTCATGTTTGAAGAGAAATGATCCAGATGGACCTCATTC 452  
Db 187 GGCTAAGGCACCATGATTTTCATGTTTGAAGAGAAATGATCCAGATGGACCTCATTC 246  
453 AGACAGATCTATTACCTTGCCTGCTGCTGAGAAACAGAAATATACATGCTTTTATTC 512  
Db 247 AGACAGATCTATTACCTTGCCTGCTGCTGAGAAACAGAAATATACATGCTTTTATTC 306  
513 TGAATTTCCCAAAACTATCAATAGCAGCAGCTTAAATGCTCTCTTGGAGCCCTCTTTT 572  
Db 307 TGAATTTCCCAAAACTATCAATAGCAGCAGCTTAAATGCTCTCTTGGAGCCCTCTTTT 366  
573 GGATCTTTTTCAGGCAACATCGGACTATGAAATGTTTCTCGAGAGAGAAATATTAAG 632  
Db 367 GGATCTTTTTCAGGCAACATCGGACTATGAAATGTTTCTCGAGAGAGAAATATTAAG 426  
633 AGAAGAAACGCAATTTGGAACAGTCGGAATTTGCTTCTTACGATTTATCACAAGGAAGTGG 692  
Db 427 AGAAGAAACGCAATTTGGAACAGTCGGAATTTGCTTCTTACGATTTATCACAAGGAAGTGG 486  
693 AACATTTCTGTTTCAAGCCGCTAGTGAATTTATACATACGTAAGAGATGGAGGCCACAAGG 752  
Db 487 AACATTTCTGTTTCAAGCCGCTAGTGAATTTATACATACGTAAGAGATGGAGGCCACAAGG 546  
753 ATTTACGCAACCACTTTAAGGCCCAATTTAGTGGAACTAGTTGTCCTCCAAATACGAT 812  
Db 547 ATTTACGCAACCACTTTAAGGCCCAATTTAGTGGAACTAGTTGTCCTCCAAATACGAT 606  
813 GGATCCCAAAATTTAGCCCGCTGATCCAACTGGATTTGCTTTATACATAGCAACGATAT 872  
Db 607 GGATCCCAAAATTTAGCCCGCTGATCCAACTGGATTTGCTTTATACATAGCAACGATAT 666  
873 TTGGATATCTAACTGCTAACACAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCT 932  
Db 667 TTGGATATCTAACTGCTAACACAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCT 726  
933 AGCCAAACATGGAAGAAGATGCCAGATCACTGGAGTCGCTACCTTTGTTCTCCAAAGAAGA 992  
Db 727 AGCCAAACATGGAAGAAGATGCCAGATCACTGGAGTCGCTACCTTTGTTCTCCAAAGAAGA 786  
993 ATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAAACAATCCAGTGGTGG 1052  
Db 787 ATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAAACAATCCAGTGGTGG 846  
1053 TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGT 1112  
Db 847 TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGT 906  
1113 TACATCCCTCTATTGTTGAACAGAGGCGCAGATTCATTTCCGTTATCTTAAACAGGTAC 1172  
Db 907 TACATCCCTCTATTGTTGAACAGAGGCGCAGATTCATTTCCGTTATCTTAAACAGGTAC 966  
1173 AGCAAAATCCTAAAGTCACCTTTTAAGATGTCCAGAAATATGATTTGATGCTGGAAGGAAGAT 1232

```
Db 967 AGCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATATATGATTTGATGCTGAAGAGGAT 1026
Qy 1233 CATAGATGTCATAGATAGGAAGTAAATTCACCTTTTGAGATCTTATTTGAAGAGATTGA 1292
Db 1027 CATAGATGTCATAGATAGGAAGTAAATTCACCTTTTGAGATCTTATTTGAAGAGATTGA 1086
Qy 1293 ATATATTGCGAGAGCTGGATGGAATCCTCTGAGGAGAAATATGCTTGGTCCATCTCTACATAGA 1352
Db 1087 ATATATTGCGAGAGCTGGATGGAATCCTCTGAGGAGAAATATGCTTGGTCCATCTCTACATAGA 1446
Qy 1353 TGGCTCCAGAGCTGCCCTACAGATAGTGTGATCTCACTGCAATTAATTTATCTCCAGATAGA 1412
Db 1147 TGGCTCCAGAGCTGCCCTGAGATAGTGTGATCTCACTGCAATTAATTTATCTCCAGATAGA 1206
Qy 1413 AGATGATGTTATGGAAGGAGAGAGCTCAATTCAGTCAGTCGCTGATCTTGTGAGGCCACT 1472
Db 1207 AGATGATGTTATGGAAGGAGAGAGCTCAATTCAGTCAGTCGCTGATCTTGTGAGGCCACT 1266
Qy 1473 AATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCAATGACATCTTTTCATGTTTT 1532
Db 1267 AATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCAATGACATCTTTTCATGTTTT 1326
Qy 1533 TCCCAAAAGTCACGAAGAGGAAATTCAGTTTTATTTTGGCTCTGAAATGCAAAAACAGGTTT 1592
Db 1327 TCCCAAAAGTCACGAAGAGGAAATTCAGTTTTATTTTGGCTCTGAAATGCAAAAACAGGTTT 1386
Qy 1593 CCGTCATTTATACAAAATACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGG 1652
Db 1387 CCGTCATTTATACAAAATACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGG 1446
Qy 1653 TGGCTGCCCTCCCAAGTGATTTCAAGTGTCTCTCAAGAGGAGATAGCAATATCCAG 1712
Db 1447 TGGCTGCCCTCCCAAGTGATTTCAAGTGTCTCTCAAGAGGAGATAGCAATATCCAG 1506
Qy 1713 TGGTGAATGGGAAGTCTTCCGCGGATGGATCTAATATCCAGTTGATGAAGTCAGAAG 1772
Db 1507 TGGTGAATGGGAAGTCTTCCGCGGATGGATCTAATATCCAGTTGATGAAGTCAGAAG 1566
Qy 1773 GCTGGTATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAG 1832
Db 1567 GCTGGTATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAG 1626
Qy 1833 TTACGTAAATCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTG 1892
Db 1627 TTACGTAAATCTGGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTG 1686
Qy 1893 CATCAGTCAGACATGTGACTCTTTTATAAGTATAGTAAACCAAGAAATCCACACTG 1952
Db 1687 CATCAGTCAGACATGTGACTCTTTTATAAGTATAGTAAACCAAGAAATCCACACTG 1746
Qy 1953 TGTGTCCCTTTACAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAACAAAGGAATT 2012
Db 1747 TGTGTCCCTTTACAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAACAAAGGAATT 1806
Qy 2013 TTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAAATTTT 2072
Db 1807 TTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAAATTTT 1866
Qy 2073 CTCCTTTGAAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCCTCATGATCT 2132
Db 1867 CTCCTTTGAAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCCTCATGATCT 1926
Qy 2133 ACAGCTGGAAAGAAATATCCTACTGTCTGCTTATATATATGTTGCTCTCAGGTGCAAGTT 2192
Db 1927 ACAGCTGGAAAGAAATATCCTACTGTCTGCTTATATATGTTGCTCTCAGGTGCAAGTT 1986
Qy 2193 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGG 2252
Db 1987 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGG 2046
Qy 2253 TTATGCTGGTTGATGACACACAGGGGATCTGTGTCACGAGGGCTTAAATTTGAAGG 2312
```

```
Db 2047 TTATGCTGGTTGATGATAGACAAACAGGGGATCTCTGTACACGAGGGCTTAAATTTGAAGG 2106
Qy 2313 GCGCTTTAAATATATAAATGGGTCAAAATAGAAATGAGCGATCAGGTGGAGGACTCCCAATA 2372
Db 2107 GCGCTTTAAATATATAAATGGGTCAAAATAGAAATGAGCGATCAGGTGGAGGACTCCCAATA 2166
Qy 2373 TCTAGCTTCTCGATATGATTTTCACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTA 2432
Db 2167 TCTAGCTTCTCGATATGATTTTCACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTA 2226
Qy 2433 TGGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTTCAAGGTTGCTAT 2492
Db 2227 TGGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTTCAAGGTTGCTAT 2286
Qy 2493 TGTGGGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTTATAT 2552
Db 2287 TGTGGGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTTATAT 2346
Qy 2553 GGGTCACCCCTGACACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA 2612
Db 2347 GGGTCACCCCTGACACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA 2406
Qy 2613 AAAGTTCCTCTGAAACCAAAATCGTTTACTGTCTTTTACATGTTTCTTGGATGAGAATGT 2672
Db 2407 AAAGTTCCTCTGAAACCAAAATCGTTTACTGTCTTTTACATGTTTCTTGGATGAGAATGT 2466
Qy 2673 CCATTTTGCACATACACAGTATATTTACTGTAGTCTTTTGTAGGGCTGGAAGCCATATGA 2732
Db 2467 CCATTTTGCACATACACAGTATATTTACTGTAGTCTTTTGTAGGGCTGGAAGCCATATGA 2526
Qy 2733 TTTACAGATCTATCTCCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTA 2792
Db 2527 TTTACAGATCTATCTCCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTA 2586
Qy 2793 TGAACCTGATCTTTTGCATCTACTCTCAAGAAACCTTGGATCAGTATTTGCTCTCTAAA 2852
Db 2587 TGAACCTGATCTTTTGCATCTACTCTCAAGAAACCTTGGATCAGTATTTGCTCTCTAAA 2646
Qy 2853 AGTGATAT 2860
Db 2647 AGTGATAT 2654

RESULT 11
AR631278 LOCUS AR631278 2671 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 2 from patent US 6844180.
ACCESSION AR631278
VERSION AR631278.1 GI:59770920
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1. (bases 1 to 2671)
AUTHORS Oi,S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 2 18-JAN-2005;
Ferring BV;;
NLX;

FEATURES
source Location/Qualifiers
1..2671
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 80.0%; Score 2495; DB 6; Length 2671;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 213 CATGCGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGGACTG 272
Db 7 CATGCGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGGACTG 66
```

QY 273 TdAGGAATATTTGAATCACAGATCGGCCTAAATTTGGAGCCTTTTTATGTTGAGCGGTA 332  
Db |||||  
QY 67 TdAGGAGATATTTGAATCACAGATCGGCCTAAATTTGGAGCCTTTTTATGTTGAGCGGTA 126  
Db |||||  
QY 333 TtCCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAGAAATATATCATGGCTACATGAT 392  
Db |||||  
QY 127 TtCCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAGAAATATATCATGGCTACATGAT 186  
Db |||||  
QY 393 GgCTAAGGACCACATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCATTC 452  
Db |||||  
QY 187 GgCTAAGGACCACATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCATTC 246  
Db |||||  
QY 453 AGACAGAACTATTTACCTTGGCCATGCTGGTGAAGAACAGAGAAATACATGTTTATTC 512  
Db |||||  
QY 247 AGACAGAACTATTTACCTTGGCCATGCTGGTGAAGAACAGAGAAATACATGTTTATTC 306  
Db |||||  
QY 513 TGAATTTCCAAAATCATCATAGAGAGCAGTCTTAATGCTCTCTTGGAGCCTCTTTT 572  
Db |||||  
QY 307 TGAATTTCCAAAATCATCATAGAGAGCAGTCTTAATGCTCTCTTGGAGCCTCTTTT 366  
Db |||||  
QY 573 GgATCTTTTTTTCAGGCAACACTGGACTATGGAAATGATTTCTCGAAGAAAGAACTATTAAAG 632  
Db |||||  
QY 367 GgATCTTTTTTTCAGGCAACACTGGACTATGGAAATGATTTCTCGAAGAAAGAACTATTAAAG 426  
Db |||||  
QY 633 AdAAAGAAACGCAITTTGGAACAGTCCGAATTCCTTTTACGATTTATCAACAAGAAAGTGG 692  
Db |||||  
QY 427 AdAAAGAAACGCAITTTGGAACAGTCCGAATTCCTTTTACGATTTATCAACAAGAAAGTGG 486  
Db |||||  
QY 693 AAcATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAAGG 752  
Db |||||  
QY 487 AAcATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAAGG 546  
Db |||||  
QY 753 ATTTAGCACAACACTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGAT 812  
Db |||||  
QY 547 ATTTAGCACAACACTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGAT 606  
Db |||||  
QY 813 GdATCCAAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACGATAT 872  
Db |||||  
QY 607 GdATCCAAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACGATAT 666  
Db |||||  
QY 873 TtGGATATCTAAACATCGTAAACGAGAGAAAGAGACTCACTTATGTCACATAGACT 932  
Db |||||  
QY 667 TtGGATATCTAAACATCGTAAACGAGAGAAAGAGACTCACTTATGTCACATAGACT 726  
Db |||||  
QY 933 AdCCAACATCGAAGAGATGCCAGATCAGCTCGAGTCGCTACCTTGTCTCCAAAGAAG 992  
Db |||||  
QY 727 AdCCAACATCGAAGAGATGCCAGATCAGCTCGAGTCGCTACCTTGTCTCCAAAGAAG 786  
Db |||||  
QY 993 AftTTGATAGATATTTGGCTATTTGGTGTCCAAAAGCTGAAAACCTCCAGTGGTG 1052  
Db |||||  
QY 787 AftTTGATAGATATTTGGCTATTTGGTGTCCAAAAGCTGAAAACCTCCAGTGGTG 846  
Db |||||  
QY 1053 TAAATTTCTTAGAATCTATAGCAAGAAATGATGAATCGAGGTGGAAATTTATCATGT 1112  
Db |||||  
QY 847 TAAATTTCTTAGAATCTATAGCAAGAAATGATGAATCGAGGTGGAAATTTATCATGT 906  
Db |||||  
QY 1113 TAcATCCCTATGTTTGAACAAGAGGCGAGATTCATTCGCTTATCCCTAAACAGAGTAC 1172  
Db |||||  
QY 907 TAcATCCCTATGTTTGAACAAGAGGCGAGATTCATTCGCTTATCCCTAAACAGAGTAC 966  
Db |||||  
QY 1173 AdCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGTCGAAGGAAGAT 1232  
Db |||||  
QY 967 AdCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGTCGAAGGAAGAT 1026  
Db |||||  
QY 1233 CATAGATGTCATAGATAAGAACTAATTTCAACTTTTGAGATTTCTTTTGAAGGATTTGA 1292  
Db |||||  
QY 1027 CATAGATGTCATAGATAAGAACTAATTTCAACTTTTGAGATTTCTTTTGAAGGATTTGA 1086  
Db |||||  
QY 1293 ATATATTTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGA 1352  
Db |||||  
QY 1087 ATATATTTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGA 1146  
Db |||||  
QY 1353 TCGETCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCCAAGTAGA 1412  
Db |||||

Db 1147 TCCTCCAGAGCTCCCTGCGAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGA 1206  
QY |||||  
QY 1413 AGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGCCTGATTTCTGTGAGCGCACT 1472  
Db |||||  
QY 1207 AGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGCCTGATTTCTGTGAGCGCACT 1266  
Db |||||  
QY 1473 AATTATCTATGAAGAAACAAAGACATCTGGATAAATATCCATGACATCTTTATGTTTT 1532  
Db |||||  
QY 1267 AATTATCTATGAAGAAACAAAGACATCTGGATAAATATCCATGACATCTTTATGTTTT 1326  
Db |||||  
QY 1533 TCCCAAGAGTCAGAGAGGAAATTTAGTTATTTTGGCCCTCTGAATCAAAAACAGGTTT 1592  
Db |||||  
QY 1327 TCCCAAGAGTCAGAGAGGAAATTTAGTTATTTTGGCTCTGGAATGAAAACAGGTTT 1386  
Db |||||  
QY 1593 CCGTCATTTATACAAAATTTACATCTATTTTAAAGSAAAGCAAAATATAAAGATCCAGTGG 1652  
Db |||||  
QY 1387 CCGTCATTTATACAAAATTTACATCTATTTTAAAGSAAAGCAAAATATAAAGATCCAGTGG 1446  
Db |||||  
QY 1653 TGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTCAAGAGGAGATAGCAATTTACCAG 1712  
Db |||||  
QY 1447 TGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTCAAGAGGAGATAGCAATTTACCAG 1506  
Db |||||  
QY 1713 TGGTGAATGGAAATTTCTTGGCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAG 1772  
Db |||||  
QY 1507 TGGTGAATGGAAATTTCTTGGCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAG 1566  
Db |||||  
QY 1773 GCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAG 1832  
Db |||||  
QY 1567 GCTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAG 1626  
Db |||||  
QY 1833 TTACGTAATCTCGAGAGAGTGACAAGGCTGACTGACCGTGGCTACTACATTTCTGCTG 1892  
Db |||||  
QY 1627 TTACGTAATCTCGAGAGAGTGACAAGGCTGACTGACCGTGGCTACTACATTTCTGCTG 1686  
Db |||||  
QY 1893 CATCAGTCACACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAATCCACACTG 1952  
Db |||||  
QY 1687 CATCAGTCACACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAATCCACACTG 1746  
Db |||||  
QY 1953 TGTGTCCTTTTACAAGCTATCAAGTCTTGAAGATGACCCAACTTGTGCAAAAACAAAGGAAT 2012  
Db |||||  
QY 1747 TGTGTCCTTTTACAAGCTATCAAGTCTTGAAGATGACCCAACTTGTGCAAAAACAAAGGAAT 1806  
Db |||||  
QY 2013 TTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATACTCTCCAGAAATTTT 2072  
Db |||||  
QY 1807 TTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATACTCTCCAGAAATTTT 1866  
Db |||||  
QY 2073 CTCTTTTGAAGTACTACTGGATTTTACATGTTATGGATGCTCTACAAAGCCTCATGATCT 2132  
Db |||||  
QY 1867 CTCTTTTGAAGTACTACTGGATTTTACATGTTATGGATGCTCTACAAAGCCTCATGATCT 1926  
Db |||||  
QY 2133 ACAGCCTGGAAGAAATATCTTACTGTGTTTCATATATGGTGGTCTCAGGTGCGAGTT 2192  
Db |||||  
QY 1927 ACAGCCTGGAAGAAATATCTTACTGTGTTTCATATATGGTGGTCTCAGGTGCGAGTT 1986  
Db |||||  
QY 2193 GGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGG 2252  
Db |||||  
QY 1987 GGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGG 2046  
Db |||||  
QY 2253 TTATGTGTTGTAGTAGACAAACAGGGGATCCTGTCAACCGAGGCTTAAATTTGAAG 2312  
Db |||||  
QY 2047 TTATGTGTTGTAGTAGACAAACAGGGGATCCTGTCAACCGAGGCTTAAATTTGAAG 2106  
Db |||||  
QY 2313 CGCCTTTAAATATAAAATGGGTCAAAATAGAAATGATGATGAGTGGAGGAGTCCCAATA 2372  
Db |||||  
QY 2107 CGCCTTTAAATATAAAATGGGTCAAAATAGAAATGATGAGTGGAGGAGTCCCAATA 2166  
Db |||||  
QY 2373 TCTAGCTTCGATATGATTTTCAATGACTTAGATCTGTGGGATCCACGGCTGGTCTTA 2432  
Db |||||  
QY 2167 TCTAGCTTCGATATGATTTTCAATGACTTAGATCTGTGGGATCCACGGCTGGTCTTA 2226  
Db |||||  
QY 2433 TGGAGGATACCTCTCCCTGATGTCATTAATGACAGAGTCCAGATATCTTTTCCAGGTTGCTAT 2492  
Db |||||



Db	2227	TGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTGAGATATCTTTCAGGGTTGCTAT	2286
Qy	2493	TGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTGATAT	2552
Db	2287	TGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTGATAT	2346
Qy	2553	GGGTCACTCTGACAGCAATGACACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA	2612
Db	2347	GGGTCACTCTGACAGCAATGACACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA	2406
Qy	2613	AAAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGCTGTTCTGATGAGATGT	2672
Db	2407	AAAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGCTGTTCTGATGAGATGT	2466
Qy	2673	CCATTTTGCATACAGATATATCTAGTCTTTTCTAGTGGGCTGAAAGCCATATGA	2732
Db	2467	CCATTTTGCATACAGATATATCTAGTCTTTTCTAGTGGGCTGAAAGCCATATGA	2526
Qy	2733	TTTACAGATCTATCTCCAGAGAGACACAGCATAGAGTTCTGATCGGAGAACATTA	2792
Db	2527	TTTACAGATCTATCTCCAGAGAGACACAGCATAGAGTTCTGATCGGAGAACATTA	2586
Qy	2793	TGAATCGATCTTTTGCATCTCTTCAAGAAACCTTGGATCAGCTATTGCTCTTAA	2852
Db	2587	TGAATCGATCTTTTGCATCTCTTCAAGAAACCTTGGATCAGCTATTGCTCTTAA	2646
Qy	2853	AGTGATAT 2860	
Db	2647	AGTGATAT 2654	
RESULT 12			
AX608725			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1			
Qi,S., Akinsanya,K.O., Riviere,P.J. and Junien,J.L.			
Novel serine protease genes related to dppiv			
Patent: WO 0231134-A 2 18-APR-2002;			
Ferring BV (NL)			
FEATURES			
source			
1..2671			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	213	CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTGCGGACTG	272
Db	7	CATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTGCGGACTG	66
Qy	273	TGAGGAAATATTGAATCAGGATCGGCCCTAAATTGGAGCCCTTTTATGTTGAGCGGTA	332
Db	67	TGAGGAAATATTGAATCAGGATCGGCCCTAAATTGGAGCCCTTTTATGTTGAGCGGTA	126
Qy	333	TTCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACCAAGAAATATCATGGCTACATGAT	392
Db	127	TTCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACCAAGAAATATCATGGCTACATGAT	186
Qy	393	GGCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC	452

Db	187	GGCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC	246
Qy	453	AGACAGAAATCTATTTACCTTCGCAATGCTCGTGAGAAACAGAGAAAATACATGTTTTATTC	512
Db	247	AGACAGAAATCTATTTACCTTCGCAATGCTCGTGAGAAACAGAGAAAATACATGTTTTATTC	306
Qy	513	TGAAATTTCCCAAATCTATCAATAGACAGAGCTTTTAATGCTCTCTTGGAAAGCCTCTTTT	572
Db	307	TGAAATTTCCCAAATCTATCAATAGACAGAGCTTTTAATGCTCTCTTGGAAAGCCTCTTTT	366
Qy	573	GGATCTTTTTCAGGCAACACTGGACTATCGAAATGTAATTTCTCGAGAAGAGAACTATTAAAG	632
Db	367	GGATCTTTTTCAGGCAACACTGGACTATCGAAATGTAATTTCTCGAGAAGAGAACTATTAAAG	426
Qy	633	AGAAAGAAAAACGCAATTTGGAAACAGTTCGTAATTTCTTACGATTTATCACCAAGGAAGTGG	692
Db	427	AGAAAGAAAAACGCAATTTGGAAACAGTTCGTAATTTCTTACGATTTATCACCAAGGAAGTGG	486
Qy	693	AACATTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACTGTAAGAGATGGAGGCCACAAAGG	752
Db	487	AACATTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACTGTAAGAGATGGAGGCCACAAAGG	546
Qy	753	ATTTTACGCAACCACTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCCCAACATACCGAT	812
Db	547	ATTTTACGCAACCACTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCCCAACATACCGAT	606
Qy	813	GGATCCAAATTTATGCCCCCTGCTGATCCAGACTGGATTCGTTTATACATAGCAACGATAT	872
Db	607	GGATCCAAATTTATGCCCCCTGCTGATCCAGACTGGATTCGTTTATACATAGCAACGATAT	666
Qy	873	TTGGATATCTAACATCGTAACACAGAGAGAGAGAGACTCACTTATGTGCAACATGAGCT	932
Db	667	TTGGATATCTAACATCGTAACACAGAGAGAGAGAGACTCACTTATGTGCAACATGAGCT	726
Qy	933	AGCCACATGAGAGAGAGACTCCAGATCAGCTGGAGTGGCTACCTTCTTCCCAAGAGA	992
Db	727	AGCCACATGAGAGAGAGACTCCAGATCAGCTGGAGTGGCTACCTTCTTCCCAAGAGA	786
Qy	993	ATTTGATAGATATCTGGCTATTTGGTGTGTCCTCCAAAGCTGAAACAACCTCCAGTGTGG	1052
Db	787	ATTTGATAGATATCTGGCTATTTGGTGTGTCCTCCAAAGCTGAAACAACCTCCAGTGTGG	846
Qy	1053	TAAATTTCTAGAAATCTATATGAAAGAAATGATGAATCTGAGGTGGAATTTATTCATGT	1112
Db	847	TAAATTTCTAGAAATCTATATGAAAGAAATGATGAATCTGAGGTGGAATTTATTCATGT	906
Qy	1113	TACATCCCTATGTTGGAAACAAAGAGGCGCAGATTCATTCGTTTATCCTAAAAACAGGTAC	1172
Db	907	TACATCCCTATGTTGGAAACAAAGAGGCGCAGATTCATTCGTTTATCCTAAAAACAGGTAC	966
Qy	1173	AGCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATTCATGCTGAAAGGAAGAT	1232
Db	967	AGCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATTCATGCTGAAAGGAAGAT	1026
Qy	1233	CATAGATGTCATAGATAAGGAACTAAATTCACCTTTTGGAGATTTCTATTTCAAGGAGTTGA	1292
Db	1027	CATAGATGTCATAGATAAGGAACTAAATTCACCTTTTGGAGATTTCTATTTCAAGGAGTTGA	1086
Qy	1293	ATATATTTGCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGA	1352
Db	1087	ATATATTTGCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGA	1146
Qy	1353	TGCTCTCCAGACTCCGCTTACAGATAGTGTTCATCTCACCTGGAATTTATTTCCAGGTAGA	1412
Db	1147	TGCTCTCCAGACTCCGCTTACAGATAGTGTTCATCTCACCTGGAATTTATTTCCAGGTAGA	1206
Qy	1413	AGATGATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTCCCTGATTTCTGTGACGCCACT	1472
Db	1207	AGATGATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTCCCTGATTTCTGTGACGCCACT	1266
Qy	1473	AATTTATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1532
Db	1267	AATTTATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1326

1533 TCCCAAAGTCACGAAGGAAATTGAGTTTATTTTTTGGCTCTGAATGCAAAACAGGTTT 1592  
Db TCCCAAAGTCACGAAGGAAATTGAGTTTATTTTTTGGCTCTGAATGCAAAACAGGTTT 1386  
1593 CGGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGG 1652  
Db CGGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGG 1446  
1653 TGGGCTGCCTGCTCCAAAGTGATTTTCAAGTGTCCTATCAAAAGAGGAGATAGCAATTAACCG 1712  
Db TGGGCTGCCTGCTCCAAAGTGATTTTCAAGTGTCCTATCAAAAGAGGAGATAGCAATTAACCG 1506  
1713 TGGTGAATCGGAAGTTCTTGGCCGCATGGAATCTAATATCAAGTTGATGAAGTCAGAG 1772  
Db TGGTGAATCGGAAGTTCTTGGCCGCATGGAATCTAATATCAAGTTGATGAAGTCAGAG 1566  
1773 GTGGTATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAG 1832  
Db GTGGTATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAG 1626  
1833 TTACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAATTTCTGCTG 1892  
Db TTACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAATTTCTGCTG 1686  
1893 CATCAGTCAGCACTGTGACTCTTTATTAAGTAAGTATAGTAACCAAGGAATCCACACTG 1952  
Db CATCAGTCAGCACTGTGACTCTTTATTAAGTAAGTATAGTAACCAAGGAATCCACACTG 1746  
1953 TGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGGAAT 2012  
Db TGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGGAAT 1806  
2013 TTGGGCACCAATTTGGATTCAGAGTCTCTCTCTGCACTACTACTCTCCAGAAATTTT 2072  
Db TTGGGCACCAATTTGGATTCAGAGTCTCTCTCTGCACTACTACTCTCCAGAAATTTT 1866  
2073 CTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAGGCTCATGATCT 2132  
Db CTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAGGCTCATGATCT 1926  
2133 ACAGCTTGGAAAGAAATATCTTACTGTCTGTCTTATATATATGGTGGTCTCAGGTGCAGTT 2192  
Db ACAGCTTGGAAAGAAATATCTTACTGTCTGTCTTATATATATGGTGGTCTCAGGTGCAGTT 1986  
2193 GGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAG 2252  
Db GGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAG 2046  
2253 TTATGTGGTTGTAGTATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAG 2312  
Db TTATGTGGTTGTAGTATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAG 2106  
2313 CGCTTTAAATATAAAATGGGTCAATAGAAATGACGATCAGGTGGAGGACTCCCAATA 2372  
Db CGCTTTAAATATAAAATGGGTCAATAGAAATGACGATCAGGTGGAGGACTCCCAATA 2166  
2373 TCTAGCTTCTCGATATGATTTCAATGACCTTACATGCTGTGGGATCCACGGCTGGTCTTA 2432  
Db TCTAGCTTCTCGATATGATTTCAATGACCTTACATGCTGTGGGATCCACGGCTGGTCTTA 2226  
2433 TGGAGGATACCTCTCCTGATGGCATTAAATGACAGAGGTGAGATCTTTCAGGGTTGCTAT 2492  
Db TGGAGGATACCTCTCCTGATGGCATTAAATGACAGAGGTGAGATCTTTCAGGGTTGCTAT 2286  
2493 TGTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTTATAT 2552  
Db TGTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTTATAT 2346  
2553 GGGTCAACCTGACAGAAAGCAAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGA 2612  
Db GGGTCAACCTGACAGAAAGCAAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGA 2406

2613 AAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGATGAGAAATGT 2672  
Db AAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGATGAGAAATGT 2466  
2673 CCATTTTGACATACACAGTATATTACTGAGTTTCTTGTAGTGGGCTGGAAGCCATATGA 2732  
Db CCATTTTGACATACACAGTATATTACTGAGTTTCTTGTAGTGGGCTGGAAGCCATATGA 2526  
2733 TTTACAGATCTATCTCAGGAGAGACACAGCATAAAGAGTTCTTGAATCGGAGAACATTA 2792  
Db TTTACAGATCTATCTCAGGAGAGACACAGCATAAAGAGTTCTTGAATCGGAGAACATTA 2586  
2793 TGAACGTCACTCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACTATTCCTGCTCTAAA 2852  
Db TGAACGTCACTCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACTATTCCTGCTCTAAA 2646  
2853 AGTGATAT 2860  
Db AGTGATAT 2654  
RESULT 13  
AR448400  
LOCUS 2797 bp DNA linear PAT 20-FEB-2004  
DEFINITION Sequence 1103 from patent US 6673549.  
ACCESSION AR448400  
VERSION AR448400.1 GI:42676724  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2797)  
AUTHORS Furness,L.M. and Buchbinder,J.I.  
TITLE Genes expressed in C3A liver cell cultures treated with steroids  
JOURNAL Patent: US 6673549-A 1103 06-JAN-2004;  
Incyte Corporation; Palo Alto, CA  
FEATURES  
source 1. .2797  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 74.2%; Score 2315; DB 6; Length 2797;  
Best Local Similarity 99.8%; Pred. No: 0;  
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;  
335 CCTGGAGTCAGCTTAAAAAGCTGCTTGCAGATACAGAAAATATCATGGCTACATGATGG 394  
Db 1 CCTGGAGTCAGCTTAAAAAGCTGCTTGCAGATACAGAAAATATCATGGCTACATGATGG 60  
395 CTAAGGCACCAATGATTTTCATGTTGTGAAGGAATGATCCAGATGCCTCATTCAG 454  
Db 61 CTAAGGCACCAATGATTTTCATGTTGTGAAGGAATGATCCAGATGCCTCATTCAG 120  
455 ACAGAACTATTACCTTGCATGCTGTGAGAAACAGAGAAATACTGCTTTTATTTCTG 514  
Db 121 ACAGAACTATTACCTTGCATGCTGTGAGAAACAGAGAAATACTGCTTTTATTTCTG 180  
515 AAATTTCCCAAACTATCAATAGAGCAGCTTAAATGCTCTTTGGAAGCTCTTTTGG 574  
Db 181 AAATTTCCCAAACTATCAATAGAGCAGCTTAAATGCTCTTTGGAAGCTCTTTTGG 240  
575 ATCTTTTTCAGGCAACATGGAATGGAATGATTTCTCGAGAGAACTATTAAGAG 634  
Db 241 ATCTTTTTCAGGCAACATGGAATGGAATGATTTCTCGAGAGAACTATTAAGAG 300  
635 AAAGAAACCGCATTTGNAAGTCGGATTCCTTCTTACGATTTACCAAGGAAGTGAA 694  
Db 301 AAAGAAACCGCATTTGNAAGTCGGATTCCTTCTTACGATTTACCAAGGAAGTGAA 360  
695 CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGCTAAAAGATGGAGGGCCACAGGAT 754  
Db 361 CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGCTAAAAGATGGAGGGCCACAGGAT 420



QY 755 TTACGCCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTGCCAACATACGGATGG 814  
Db 421 TTACGCCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTGCCAACATACGGATGG 480  
QY 815 ATCCAAAATTAATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 874  
Db 481 ATCCAAAATTAATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 540  
QY 875 GGATATCTAACTGTAACCCAGAGAAAGAGAGACTCACATTATGTGCACAAATGAGCTAG 934  
Db 541 GGATATCTAACTGTAACCCAGAGAAAGAGAGACTCACATTATGTGCACAAATGAGCTAG 600  
QY 935 CCAACATGGAAGAGATGCCAGATCAGCTGGATCGCTACCTTTGTTCTCCAGAGAAAT 994  
Db 601 CCAACATGGAAGAGATGCCAGATCAGCTGGATCGCTACCTTTGTTCTCCAGAGAAAT 660  
QY 995 TTGATAGATATCTGCTATTCGTTGTCGAAAGCTGAAACAACTCCCAAGTGGTGA 1054  
Db 661 TTGATAGATATCTGCTATTCGTTGTCGAAAGCTGAAACAACTCCCAAGTGGTGA 720  
QY 1055 AATATCTTGAATCTTATATGAAGAAATGATGATCTGAGGTGGAAATTAATCATGTTA 1114  
Db 721 AATATCTTGAATCTTATATGAAGAAATGATGATCTGAGGTGGAAATTAATCATGTTA 780  
QY 1115 CATCCCTATGTTGGAACAAGG-AGGCGAGATTCAATTCGTTATCTCTAAAAACAGGTACA 1173  
Db 781 CATCCCTATGTTGGAACAAGGCGAGAGATTCAATTCGTTATCTCTAAAAACAGGTACA 840  
QY 1174 GCAATCTTAAGTCACTTTTAAAGATGTCAAGAAATAATGATGCTGGAAGGAAGATC 1233  
Db 841 GCAATCTTAAGTCACTTTTAAAGATGTCAAGAAATAATGATGCTGGAAGGAAGATC 900  
QY 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCTTATTTGAAGGATTGAA 1293  
Db 901 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCTTATTTGAAGGATTGAA 960  
QY 1294 TATATTTGCCAGAGCTGGATGGACTCCTGAGGGAATAA---TGCTTGGTCCATCTACT 1349  
Db 961 TATATTTGCCAGAGCTGGATGGACTCCTGAGGGAATAA---TGCTTGGTCCATCTACT 1020  
QY 1350 AGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGGAATTTATTTCCCACT 1409  
Db 1021 AGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGGAATTTATTTCCCACT 1080  
QY 1410 AGAAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCACTGCTGATCTGTGACGCC 1469  
Db 1081 AGAAGATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCACTGCTGATCTGTGACGCC 1140  
QY 1470 ACTAATTTATGAGAGAAACAACAGACATCTGGATTAATATCCATGACATCTTTTCATGT 1529  
Db 1141 ACTAATTTATGAGAGAAACAACAGACATCTGGATTAATATCCATGACATCTTTTCATGT 1200  
QY 1530 TTTTCCCAAGTCAAGAGGAAATTTGATTTATTTTGGCTCTGAAATGCAAAACAGG 1589  
Db 1201 TTTTCCCAAGTCAAGAGGAAATTTGATTTATTTTGGCTCTGAAATGCAAAACAGG 1260  
QY 1590 TTTCCGTCATTTATACAAATTAATCTATTTTAAAGGAAGCAAAATATAAACGATCCAG 1649  
Db 1261 TTTCCGTCATTTATACAAATTTATCTATTTTAAAGGAAGCAAAATATAAACGATCCAG 1320  
QY 1650 TGGTGGGCTGCTGCTCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTAAC 1709  
Db 1321 TGGTGGGCTGCTGCTCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTAAC 1380  
QY 1710 CAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAG 1769  
Db 1381 CAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAG 1440  
QY 1770 AAGGCTGGTATTTTGAAGGCAACCAAGACTCCCTTTTAGAGCATCACTGTGAGTAGT 1829  
Db 1441 AAGGCTGGTATTTTGAAGGCAACCAAGACTCCCTTTTAGAGCATCACTGTGAGTAGT 1500

QY 1830 CAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCGTCTACTACATTCTTTG 1889  
Db 1501 CAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCGTCTACTACATTCTTTG 1560  
QY 1890 CTGCATCAGTCAAGCATCTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACA 1949  
Db 1561 CTGCATCAGTCAAGCATCTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACA 1620  
QY 1950 CTGTGTCTCCCTTTACAGCTATCAAGTCTGAAATGATGACCAACTTCCAAAAACAAAGGA 2009  
Db 1621 CTGTGTCTCCCTTTACAGCTATCAAGTCTGAAATGATGACCAACTTCCAAAAACAAAGGA 1680  
QY 2010 ATTTTGGGCCACCATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAAT 2069  
Db 1681 ATTTTGGGCCACCATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAAT 1740  
QY 2070 TTTCTCTTTTGAAGTACTACTGGATTTTACATTGTATGGGATGCTCTCAAGGCTCATGA 2129  
Db 1741 TTTCTCTTTTGAAGTACTACTGGATTTTACATTGTATGGGATGCTCTCAAGGCTCATGA 1800  
QY 2130 TCTACAGCCTGGAAAGAAATATCTACTGTGTCTCATATATATGTGTGCTCTCAAGGTGCA 2189  
Db 1801 TCTACAGCCTGGAAAGAAATATCTACTGTGTCTCATATATATGTGTGCTCTCAAGGTGCA 1860  
QY 2190 GTTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCT 2249  
Db 1861 GTTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCT 1920  
QY 2250 AGGTTATGTGGTTGTAGTAGACACAACAGGGGATCTGTCCACCGGGGCTTAAATTTGA 2309  
Db 1921 AGGTTATGTGGTTGTAGTAGACACAACAGGGGATCTGTCCACCGGGGCTTAAATTTGA 1980  
QY 2310 AGCGGCTTTTAAATATAATTAAGTGGTCAAAATAGAAATTTGACGATCAGTGGAGACTCCA 2369  
Db 1981 AGCGGCTTTTAAATATAATTAAGTGGTCAAAATAGAAATTTGACGATCAGTGGAGACTCCA 2040  
QY 2370 ATATCTAGCTTCTCCATATGATTTTCATTGACTTATAGTCTGTGTGGGCATCCACGGCTGGTC 2429  
Db 2041 ATATCTAGCTTCTCCATATGATTTTCATTGACTTATAGTCTGTGTGGGCATCCACGGCTGGTC 2100  
QY 2430 CTATGAGGATACCTCTCCCTGATGGCAATTAATCGAGAGTCAAGATATCTTTCAGGCTTGC 2489  
Db 2101 CTATGAGGATACCTCTCCCTGATGGCAATTAATCGAGAGTCAAGATATCTTTCAGGCTTGC 2160  
QY 2490 TATTTGCTGGGGCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTA 2549  
Db 2161 TATTTGCTGGGGCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTA 2220  
QY 2550 TATGGGTCACCTGACCCAGAAATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGC 2609  
Db 2221 TATGGGTCACCTGACCCAGAAATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGC 2280  
QY 2610 AGAAAAGTTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAA 2669  
Db 2281 AGAAAAGTTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAA 2340  
QY 2670 TGTCCATTTTGCACATACAGTATATTACTGAGTTTTTTAGTCAGGCTGGAAAGCCATA 2729  
Db 2341 TGTCCATTTTGCACATACAGTATATTACTGAGTTTTTTAGTCAGGCTGGAAAGCCATA 2400  
QY 2730 TCAATTTACAGATCTATCTCAGGAGACACACATAGAGTTCTCGAATCCGGGAGAA 2789  
Db 2401 TCAATTTACAGATCTATCTCAGGAGACACACATAGAGTTCTCGAATCCGGGAGAA 2460  
QY 2790 TTATGAACTGCATCTTTTGGACTTACCTTCAAGAAACCTTGGATACAGTATTTGCTGCTCT 2849  
Db 2461 TTATGAACTGCATCTTTTGGACTTACCTTCAAGAAACCTTGGATACAGTATTTGCTGCTCT 2520  
QY 2850 AAAAGTGAATTAATTTTGAACCTGTGTAGAACTCTCTGGTATACATCTGGGCTATTTAACCAA 2909  
Db 2521 AAAAGTGAATTAATTTTGAACCTGTGTAGAACTCTCTGGTATACATCTGGGCTATTTAACCAA 2580  
QY 2910 ATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATATTTTGTATCTGCTCAT 2969

```
Db 2581 ATGAGGAGTTTAAATCAACAGAAAACAGAAATTGATCATCACATTTTGATACCTGCCAT 2640
Qy 2970 GTAACATCTACTCTGAAATAAATCTGTGTCATCGAGGGTCTTACGGTTTGTGTAGT 3029
Db 2641 GTAAATCTACTCTGAAATAAATCTGTGTCATCGAGGGTCTACGGTTTGTGTAGT 2700
Qy 3030 AATCTAATACCTTAACCCCATCTCAAAATCAAAATGATACATATTCCTGAGAGCCCA 3089
Db 2701 AATCTAATACCTTAACCCCATCTCAAAATCAAAATGATACATATTCCTGAGAGCCCA 2760
Qy 3090 GCAATACATAAGAAATTAATAAAAAAAA 3120
Db 2761 GCAATACATAAGAAATTAATAAAAAAAA 2791

RESULT 14
AR631281 4523 bp DNA linear PAT 14-FEB-2005
LOCUS Sequence 8 from patent US 6844180.
DEFINITION AR631281
ACCESSION AR631281
VERSION AR631281.1 GI:59770924
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4523)
AUTHORS O.I.S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 8 JAN-2005;
Ferring BV;;
NLX; Location/Qualifiers
1..4523
/organism="unknown"
/mol_type="genomic DNA"

FEATURES
source
ORIGIN
Query Match 68.2%; Score 2128; DB 6; Length 4523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTGTCTAAGCTTCGAGGCCAAGCCGCTGCTACTGCCGCCGCTGCTTCTTAGTGCCG 60
Db 1 AGTGTCTAAGCTTCGAGGCCAAGCCGCTGCTACTGCCGCCGCTGCTTCTTAGTGCCG 60
Qy 61 CGTTCGCCGCTGGGTTGTACCGCGCGCGCGCGAGGAGCCACTGCAACAGACCG 120
Db 61 CGTTCGCCGCTGGGTTGTACCGCGCGCGCGCGAGGAGCCACTGCAACAGACCG 120
Qy 121 GAGTGGAGGCGCGCAGCATGAAGCGCGCGCGCGCTCATAGCCAGCTCGGACCG 180
Db 121 GAGTGGAGGCGCGCAGCATGAAGCGCGCGCGCGCTCATAGCCAGCTCGGACCG 180
Qy 181 TCGCGCGCGCGCGCGGAGAAATGCAATGCAATGCGCAGCAGCAATGGAACAGAACAG 240
Db 181 TCGCGCGCGCGCGCGGAGAAATGCAATGCAATGCGCAGCAGCAATGGAACAGAACAG 240
Qy 241 CTGGGTCTTCAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Db 241 CTGGGTCTTCAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300
Qy 301 CCTAAATGGAGCCTTTTATGTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360
Db 301 CCTAAATGGAGCCTTTTATGTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360
Qy 361 GCCGATACAGAAAATATCATGTCTACATGATGGCTTAAGGACCAACATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAAATATCATGTCTACATGATGGCTTAAGGACCAACATGATTTTCATGTTT 420
Qy 421 GTGAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAAATCTATTACCTTGGCATGTCT 480
Db 421 GTGAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAAATCTATTACCTTGGCATGTCT 480
```

```
481 GGTGAGAACAGAGAAAATACACTGTTTATCTGAAATTTCCAAAACTATCAATAGACCA 540
481 GGTGAGAACAGAGAAAATACACTGTTTATCTGAAATTTCCAAAACTATCAATAGACCA 540
541 GCAGTCTTAATGCTCTCTTGGAAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGACTAT 600
541 GCAGTCTTAATGCTCTCTTGGAAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGACTAT 600
601 GGAATGTAATCTCGAGAAAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
601 GGAATGTAATCTCGAGAAAGAACTATTAAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
661 ATTGCTTCTTACGATTTATCAACGAAGAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA 720
661 ATTGCTTCTTACGATTTATCAACGAAGAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA 720
721 ATTTATACGTAAGAGTGGAGGCCACAGAGATTTTACGCAACAACTTTTAAGGCCCAAT 780
721 ATTTATACGTAAGAGTGGAGGCCACAGAGATTTTACGCAACAACTTTTAAGGCCCAAT 780
781 CTAGTGGAACTAGTGTCTCCCAACATACGATGGATCCAAATTTATGCCCCCTGATCCA 840
781 CTAGTGGAACTAGTGTCTCCCAACATACGATGGATCCAAATTTATGCCCCCTGATCCA 840
841 GACTGGATTGCTTTTATACATAGCAACGATTTTGGATATCTTAACATCTGAACCAAGAA 900
841 GACTGGATTGCTTTTATACATAGCAACGATTTTGGATATCTTAACATCTGAACCAAGAA 900
901 GAAAGAGACTCATTATGTGCAAAAGTACGCAACATGCAAGAGTGAAGAGTGCAGATCA 960
901 GAAAGAGACTCATTATGTGCAAAAGTACGCAACATGCAAGAGTGAAGAGTGCAGATCA 960
961 GCTGGAGTGCCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
961 GCTGGAGTGCCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
1021 TGTCCAAAAGCTGAAACCAACTCCAGTGTGTGTAATAATTTCTAGATATTTCTGGCTATTGGTGG 1080
1021 TGTCCAAAAGCTGAAACCAACTCCAGTGTGTGTAATAATTTCTAGATATTTCTGGCTATTGGTGG 1080
1081 AATGATGATCTGAGTGTGTAATAATTTCTAGATATTTCTGGCTATTGGTGG 1140
1081 AATGATGATCTGAGTGTGTAATAATTTCTAGATATTTCTGGCTATTGGTGG 1140
1141 GCAGATTCAATCCGTTATCTTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAGATG 1200
1141 GCAGATTCAATCCGTTATCTTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAGATG 1200
1201 TCAGAAATAATGATTTGATGCTGAAGAAAGGATCATAGATGTCTAGATTAAGGAACCTAAT 1260
1201 TCAGAAATAATGATTTGATGCTGAAGAAAGGATCATAGATGTCTAGATTAAGGAACCTAAT 1260
1261 CAACCTTTTGAAGATTTCTTTTGAAGAGTGTGAATATTTTGCAGAGCTGGAATGCACTCT 1320
1261 CAACCTTTTGAAGATTTCTTTTGAAGAGTGTGAATATTTTGCAGAGCTGGAATGCACTCT 1320
1321 GAGGAAAATATATGCTTTGGTCCATCTCTAGATAGATCCCTCCAGACTCGCTTACATAGTG 1380
1321 GAGGAAAATATATGCTTTGGTCCATCTCTAGATAGATCCCTCCAGACTCGCTTACATAGTG 1380
1381 TTGATCTCACTGAAATTTATTTTCCAGTGAAGATGATTTTGAAGAGGAGGAGGAGCTC 1440
1381 TTGATCTCACTGAAATTTATTTTCCAGTGAAGATGATTTTGAAGAGGAGGAGGAGCTC 1440
1441 ATTGAGTCACTGCTGATTTCTGTGAGCCCACTAATTTATCTATGAAGAAACCAACAGATC 1500
1441 ATTGAGTCACTGCTGATTTCTGTGAGCCCACTAATTTATCTATGAAGAAACCAACAGATC 1500
1501 TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCAAGAGAGGAAATGAG 1560
1501 TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCAAGAGAGGAAATGAG 1560
```

Qy	1561	TTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCAATTTATACAAAATTACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCAATTTATACAAAATTACATCTATT	1620
Qy	1621	TTAAAGGAAAGCAAAATAAAACGATCCAGTGGTGGGCTGCCTCCACCAAGTATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAAAATAAAACGATCCAGTGGTGGGCTGCCTCCACCAAGTATTTCAAG	1680
Qy	1681	TGTCCTATCAAGAGGAGATAGCAATTAACCAAGTGGTGAATGGGAGTTCTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAGAGGAGATAGCAATTAACCAAGTGGTGAATGGGAGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAAATATCCAGTTGATGAGTCAAGTCAAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Db	1741	GGATCTAAATATCCAGTTGATGAGTCAAGTCAAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Qy	1801	TCCCTCTTAGAGCATCACTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTCAACAAG	1860
Db	1801	TCCCTCTTAGAGCATCACTGTAGTGTAGTCAAGTAAATCCTGGAGAGGTCAACAAG	1860
Qy	1861	CTGACTGACCGTGGTACTCACTTCTGTGTCATCAGTCAGCACTGTGACTCTCTTTATA	1920
Db	1861	CTGACTGACCGTGGTACTCACTTCTGTGTCATCAGTCAGCACTGTGACTCTCTTTATA	1920
Qy	1921	AGTAAGTATAGTAACCAAGAATCCACACTGTGTGTCCCTTTTACAGCTATCAAGTCCT	1980
Db	1921	AGTAAGTATAGTAACCAAGAATCCACACTGTGTGTCCCTTTTACAGCTATCAAGTCCT	1980
Qy	1981	GAAGATGACCAACTTSCAAAACAAAGGAATTTTGGGCAACATTTTGGATTCAGCAGGT	2040
Db	1981	GAAGATGACCCAACTTTCGAAAACAAAGGAATTTTGGGCCAACATTTTGGATTCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTATACCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTACA	2100
Db	2041	CCTCTTCTGACTATACCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTACA	2100
Qy	2101	TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAAGAAATATCCTACTGTG	2160
Db	2101	TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAAGAAATATCCTACTGTG	2160
Qy	2161	CTGTTTCATATGGTGTC	2179
Db	2161	CTGTTTCATATGGTGTC	2179

RESULT 15					
AX608731					
LOCUS	AX608731	4523 bp	DNA	linear	PAT 17-FEB-2003
DEFINITION	Sequence 8 from Patent WO02311134.				

SOURCE	ORGANISM
Homo sapiens (human)	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.	Homo sapiens

1 Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.  
 2 Novel serine protease genes related to dppiv  
 3 Patent: WO 03/1134-A 8 18-APR-2002;  
 4  
 5

```

ORIGIN
Query Match 60.2% Score 2128; DB 6; Length 4523;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
/db_xref=taxon:9606"

```

Qy	1	AAGTGTAAAGCTCCGAGGCCAAGGCCGCTGCTACTGCCGCGCTGCTTCTTAGTGCCG	60
Db	1	AAGTGTAAAGCTCCGAGGCCAAGGCCGCTGCTACTGCCGCGCTGCTTCTTAGTGCCG	60
Qy	61	CGTTCCGCGCTGGGTTGTACCCGCGCGCCGCGGAGAACCCACTGCACACGAGACCG	120
Db	61	CGTTCCGCGCTGGGTTGTACCCGCGCGCCGCGGAGAACCCACTGCACACGAGACCG	120
Qy	121	GAGTGAGCGCGCGCAGCATGAACGCGCGCAGGCCCGCTCCATAGACGACGCTCGGACCG	180
Db	121	GAGTGAGCGCGCGCAGCATGAACGCGCGCAGGCCCGCTCCATAGACGACGCTCGGACCG	180
Qy	181	TCGCGCGCGCGCGGGGGAAGAAAATGCAACATGGCAGCAGCAGCAATGGAAACAGAACAG	240
Db	181	TCGCGCGCGCGCGGGGGAAGAAAATGCAACATGGCAGCAGCAGCAGCAATGGAAACAGAACAG	240
Qy	241	CTGGGCTGTTGAGATATTTGAAATCTCGGACTGTGAGGAGATNTTGAATCACAGGATCGG	300
Db	241	CTGGGCTGTTGAGATATTTGAAATCTCGGACTGTGAGGAGATNTTGAATCACAGGATCGG	300
Qy	301	CCTAAATTTGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT	360
Db	301	CCTAAATTTGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT	360
Qy	361	GCCGATACCGAAGAAATCATGTGCTACATGATGGCTAAGGCAACACATGATTCATGTTT	420
Db	361	GCCGATACCGAAGAAATCATGTGCTACATGATGGCTAAGGCAACACATGATTCATGTTT	420
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCACAGACAGAAATCTATTACCTTGGCCATGTCT	480
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCACAGACAGAAATCTATTACCTTGGCCATGTCT	480
Qy	481	GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTCCTCAAACTCAATAGAGCA	540
Db	481	GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTCCTCAAACTCAATAGAGCA	540
Qy	541	CGAGCTTAAAGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGACATAT	600
Db	541	CGAGCTTAAAGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGACATAT	600
Qy	601	GGAATGPTATCTCGAAGAAAGAACTATTAAGAAAGAAAGAAACGCAATTTGGAACAGTCGGA	660
Db	601	GGAATGPTATCTCGAAGAAAGAACTATTAAGAAAGAAAGAAACGCAATTTGGAACAGTCGGA	660
Qy	661	ATTGCTCTTACGATATCACAAGGAAGTGGAACTATTTCTGTTTCAAGCGGATGTTGGA	720
Db	661	ATTGCTCTTACGATATCACAAGGAAGTGGAACTATTTCTGTTTCAAGCGGATGTTGGA	720
Qy	721	ATTATATCAGTAAAAAGATGGAGGCCACAAGGATTTTACGCAACACCTTTAAGGCCAAT	780
Db	721	ATTATATCAGTAAAAAGATGGAGGCCACAAGGATTTTACGCAACACCTTTAAGGCCAAT	780
Qy	781	CTAGTGGAAACTAGTTGTGCCACATACGATGGATCCCAAAATTTACGCGCTGATCCA	840
Db	781	CTAGTGGAAACTAGTTGTGCCACATACGATGGATCCCAAAATTTACGCGCTGATCCA	840
Qy	841	GACTGGATTGCTTTTATACATACGACAGATNTTGGATATCTAACATCGTAACACAGAA	900
Db	841	GACTGGATTGCTTTTATACATACGACAGATNTTGGATATCTAACATCGTAACACAGAA	900
Qy	901	GAAGGAGACTCACTATGTGCAATGACATGCTAGCCCAATGGAGAGATGCCAGATCA	960
Db	901	GAAGGAGACTCACTATGTGCAATGACATGCTAGCCCAATGGAGAGATGCCAGATCA	960
Qy	961	GCTGGAGTCGCTACCTTTGCTTCCAAAGAAATTTGATATTTCTGCTATTGGTGG	1020
Db	961	GCTGGAGTCGCTACCTTTGCTTCCAAAGAAATTTGATATTTCTGCTATTGGTGG	1020
Qy	1021	TGTCCAAAGCTGAACAACTCCAGTGGTGTAATTTCTTAAATTTCTATATGAAGAA	1080
Db	1021	TGTCCAAAGCTGAACAACTCCAGTGGTGTAATTTCTTAAATTTCTATATGAAGAA	1080
Qy	1081	AATGATGATCTGAGGTGGAAATTTATCATCTCCCTATGTTTGGAAACAGAGGG	1140



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 04:08:39 ; Search time 1770 Seconds  
(without alignments)  
11747.933 Million cell updates/sec

Title: US-10-825-632-2  
Perfect score: 3120  
Sequence: 1 aagtgctaaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring tables: OMIGO NUG  
Gapop 60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size: 6

Total number of hits satisfying chosen parameters: 9993364

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_21.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002as.\*
- 7: Geneseq2002bs.\*
- 8: Geneseq2003as.\*
- 9: Geneseq2003bs.\*
- 10: Geneseq2003cs.\*
- 11: Geneseq2003ds.\*
- 12: Geneseq2004as.\*
- 13: Geneseq2004bs.\*
- 14: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3120	100.0	3120	4 AAC85694	Aac85694 Nucleotid
2	3120	100.0	3120	6 AAD38956	Aad38956 Human dip
3	3003	96.2	3106	6 ABK12892	Abk12892 Human pro
4	2957	94.8	3143	6 AAH99934	Aah99934 cDNA enco
5	2889	92.6	4829	6 ABK83327	Abk83327 cDNA enco
6	2598	83.3	2649	14 ADV43981	Adv43981 Human psy
7	2541	81.4	2643	6 AAH99935	Aah99935 Coding se
8	2496	80.0	2649	6 ABX12255	Abx12255 cDNA enco
9	2495	80.0	2671	6 ABK83322	Abk83322 cDNA enco
10	2315	74.2	2797	12 ADL13374	Adl13374 Human ste
11	2305	73.9	2929	10 ACA92425	Aca92425 DNA enco
12	2271	72.8	2952	10 ACA92421	Aca92421 DNA enco
13	2217	71.1	2696	12 ADI16386	Adi16386 Human pro
14	2128	68.2	4523	6 ABK83325	Abk83325 cDNA enco
15	2128	68.2	4576	6 ABK83331	Abk83331 cDNA enco
16	2116	67.8	2842	6 ABN59774	Abn59774 Novel hum
17	1984	63.6	4685	6 ABK83332	Abk83332 cDNA enco
18	1950	62.5	2083	7 ADR41222	Adr41222 Human CD-
19	1922	61.6	2349	14 ADV43982	Adv43982 Human psy

20	1770	56.7	1821	6	ABV76411	Abv76411 Dipeptidyl
21	1450	46.5	2161	4	AAH15009	Aah15009 Human cDN
22	1377	44.1	2668	6	ABN59775	Abn59775 Novel hum
23	1191	38.2	4309	6	ABK83328	Abk83328 cDNA enco
24	1164	37.3	2510	6	AAD23843	Aad23843 Human pro
25	1023	32.8	1669	4	AAC85696	Aac85696 Nucleotid
26	956	30.6	2702	13	ADT04072	Adt04072 Human pro
27	877	28.1	1356	6	ABK83326	Abk83326 cDNA enco
28	823	26.4	2251	10	AD879035	Ad879035 Human pro
29	790	25.3	1197	4	AAC85695	Aac85695 Nucleotid
30	789	25.3	1083	4	AAC85697	Aac85697 Nucleotid
31	760	24.4	832	6	ABK83330	Abk83330 cDNA enco
32	739	23.7	873	4	AAF81719	Aaf81719 Human pro
33	737	23.6	925	6	ABL90148	AbL90148 Human pol
34	654	21.0	2463	10	ACA92424	Aca92424 DNA enco
35	501	16.1	631	4	AAH07860	Aah07860 Human cDN
36	472	15.1	620	6	ABK83329	Abk83329 cDNA enco
37	431	13.8	502	9	ACH32373	Ach32373 Human end
38	309	9.9	823	6	ABK30401	Abk30401 Human G-p
39	254	8.1	561	4	AAH00876	Aah00876 Human rep
40	205	6.6	587	4	AAH12830	Aah12830 Human cDN
41	167	5.4	168	8	ABZ09094	Abz09094 Human oli
42	167	5.4	168	10	ABZ78547	Abz78547 Tumour su
43	156	5.0	4797	4	AAH04386	Aal04386 Human rep
44	82	2.6	89	3	AAC12358	Aac12358 Human sec
45	46	1.5	60	6	ABN42337	Abn42337 Human spl

ALIGNMENTS

RESULT 1  
AAC85694  
ID AAC85694 standard; cDNA; 3120 BP.  
XX  
AC AAC85694;  
XX  
DT 29-JUN-2001 (first entry)  
XX  
DE Nucleotide sequence of human DPP8.  
XX  
KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;  
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;  
KW growth hormone deficiency; Glucose level; mucosal regeneration;  
KW non-insulin dependent diabetes mellitus; glucose intolerance;  
KW immunosuppression; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key : Location/Qualifiers  
FT CDS 214..2862  
FT /tag= a  
FT /product= "Human DPP8"  
XX

WO200119866-A1.  
22-MAR-2001.  
11-SEP-2000; 2000WO-AU001085.  
10-SEP-1999; 99AU-00002762.  
18-FEB-2000; 2000AU-00005709.  
(UNSY ) UNIV SYDNEY.  
Abbott CA, Gorell MD;  
WPI; 2001-281520/29.  
P-PSDB; AAB47187.  
New human dipeptidyl aminopeptidase (DPP8) useful for cleaving substrates, identifying inhibitors of DPP8 catalytic activity which have therapeutic uses, and for detecting activated T cells.









Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGCTGCTTCTTAGTGCCG 60  
Qy 61 CGTTCCGCGCGCTGGGTTGTACCGGCGCGCCGCCGAGGAAGCACTGCAACACGAGACCG 120  
Db 61 CGTTCCGCGCGCTGGGTTGTACCGGCGCGCCGCCGAGGAAGCACTGCAACACGAGACCG 120  
Qy 121 GAGTGGAGCGCGCGCAGCATGAGAGCGCGCAGCGCGCTCCATAGCGCACGTCGGGACGG 180  
Db 121 GAGTGGAGCGCGCGCAGCATGAGAGCGCGCAGCGCGCTCCATAGCGCACGTCGGGACGG 180  
Qy 181 TCCGCGCGCGCGCGGGGAGGAAGATGCAACATCGCAGCAGCAATGGAACACAGACAG 240  
Db 181 TCCGCGCGCGCGCGGGGAGGAAGATGCAACATCGCAGCAGCAATGGAACACAGACAG 240  
Qy 241 CTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCA CAGGATCGG 300  
Db 241 CTGGGTGTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCA CAGGATCGG 300  
Qy 301 CCTAAATTGGAGCCTTTTATGTGTGACGGGTATTCCTGGAGTCAGCTTTAAAAAGCTGCTT 360  
Db 301 CCTAAATTGGAGCCTTTTATGTGTGACGGGTATTCCTGGAGTCAGCTTTAAAAAGCTGCTT 360  
Qy 361 GCCGATACCGAAAAATATCATGTGTACATGATGGCTAAGGCCACCAACATGATTTTCATGTTT 420  
Db 361 GCCGATACCGAAAAATATCATGTGTACATGATGGCTAAGGCCACCAACATGATTTTCATGTTT 420  
Qy 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTTGCCATGTCT 480  
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTTGCCATGTCT 480  
Qy 481 GGTGAGAACAGAGAAAAATACACTGTTTTTATCTGAAATTTCCCAAACTATCAATAGAGCA 540  
Db 481 GGTGAGAACAGAGAAAAATACACTGTTTTTATCTGAAATTTCCCAAACTATCAATAGAGCA 540  
Qy 541 GCAGTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
Db 541 GCAGTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
Qy 601 GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACCAATCTGNAACGTGCGA 660  
Db 601 GGAATGTATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACCAATCTGNAACGTGCGA 660  
Qy 661 ATTGCTTTCTACGATTTATCAGCAAGGAAGTGAACATTTCTGTTTCAAGCGGTAGTGGG 720  
Db 661 ATTGCTTTCTACGATTTATCACCAGGAGTGAACATTTCTGTTTCAAGCGGTAGTGGG 720  
Qy 721 ATTTATCAGCTAAAGATGAGGGCCAAGGATTTTACGCAACAACTTTAAGGCCCAAT 780  
Db 721 ATTTATCAGCTAAAGATGAGGGCCAAGGATTTTACGCAACAACTTTAAGGCCCAAT 780  
Qy 781 CTAGTGAAACTAGTTGTGCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCA 840  
Db 781 CTAGTGAAACTAGTTGTGCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCA 840  
Qy 841 GACTGGATTCCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACACAGAGAA 900  
Db 841 GACTGGATTCCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACACAGAGAA 900  
Qy 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAAAGATGCCAGATCA 960  
Db 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAAAGATGCCAGATCA 960  
Qy 961 GCTGGAGTCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATCTGCTTATTTGGTGG 1020  
Db 961 GCTGGAGTCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATCTGCTTATTTGGTGG 1020  
Qy 1021 TGTCCAAAAGCTGAACAACTCCAGTGTGTGTAATTTCTTAGAATTTCTATATGAAGAA 1080  
Db 1021 TGTCCAAAAGCTGAACAACTCCAGTGTGTGTAATTTCTTAGAATTTCTATATGAAGAA 1080  
Qy 1081 AATGATGAACTCTGAGGTGAATATTATCTGTTACATCCCTATTTGGAAACAGGAGG 1140  
Db 1081 AATGATGAACTCTGAGGTGAATATTATCTGTTACATCCCTATTTGGAAACAGGAGG 1140

Qy 1141 GCAGATTCATTCGGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACCTTTTAGATG 1200  
Db 1141 GCAGATTCATTCGGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACCTTTTAGATG 1200  
Qy 1201 TCAGAAATATGATTTGATGCTGAAGGAAGATCATAGATGTCTATAGATAAGGAACATAAT 1260  
Db 1201 TCAGAAATATGATTTGATGCTGAAGGAAGATCATAGATGTCTATAGATAAGGAACATAAT 1260  
Qy 1261 CAACCTTTTGAGATTTCTATTTTGAAGGAGTTGAATATATTGCGAGAGCTGGATGGACTCCT 1320  
Db 1261 CAACCTTTTGAGATTTCTATTTTGAAGGAGTTGAATATATTGCGAGAGCTGGATGGACTCCT 1320  
Qy 1321 GAGGAAAAATATGCTTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380  
Db 1321 GAGGAAAAATATGCTTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380  
Qy 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC 1440  
Db 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC 1440  
Qy 1441 ATTGAGTCAGTGCCTGANTCTGTGACGCCACTAAATTTATCTATGAAGAAACACAGACATC 1500  
Db 1441 ATTGAGTCAGTGCCTGANTCTGTGACGCCACTAAATTTATCTATGAAGAAACACAGACATC 1500  
Qy 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCCAGAGAGRAATTTGAG 1560  
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCCAGAGAGRAATTTGAG 1560  
Qy 1561 TTTATTTTGGCTCTCAATGCAAAACAGGTTTCCGTCATTTTATACAAATTTTACATCTATT 1620  
Db 1561 TTTATTTTGGCTCTCAATGCAAAACAGGTTTCCGTCATTTTATACAAATTTTACATCTATT 1620  
Qy 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAG 1680  
Db 1621 TTAAGGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAG 1680  
Qy 1681 TGCTCTATCAAGAGAGATAGCAATTTACAGTGTGTGAATGGGAAGTTCTTTGGCCGGCAT 1740  
Db 1681 TGCTCTATCAAGAGAGATAGCAATTTACAGTGTGTGAATGGGAAGTTCTTTGGCCGGCAT 1740  
Qy 1741 GGATCTAATATCAAGTTGATCAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800  
Db 1741 GGATCTAATATCAAGTTGATCAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800  
Qy 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCTCGAGAGGTCACAAGG 1860  
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCTCGAGAGGTCACAAGG 1860  
Qy 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAAGCACTGTGACTTCTTTATA 1920  
Db 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAAGCACTGTGACTTCTTTATA 1920  
Qy 1921 AGTAAGTATAGTAACCCAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCCT 1980  
Db 1921 AGTAAGTATAGTAACCCAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCCT 1980  
Qy 1981 GAAGATGACCCCAACTTGGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040  
Db 1981 GAAGATGACCCCAACTTGGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040  
Qy 2041 CCTCTTCCTGACTATCTCTCCAGAAATTTCTTTTGAAGTACTACTGGATTATCA 2100  
Db 2041 CCTCTTCCTGACTATCTCTCCAGAAATTTCTTTTGAAGTACTACTGGATTATCA 2100  
Qy 2101 TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTG 2160  
Db 2101 TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTG 2160  
Qy 2161 CTGTTTCATATATGTTGGTCTCTCAGGTGAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAG 2220  
Db 2161 CTGTTTCATATATGTTGGTCTCTCAGGTGAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAG 2220

QY 2221 TATTTCCGCTTGAATACCTTAGCCTCTAGCTTATGTTGTTAGTAGATAGACACAGG 2280  
DB |||||  
QY 2221 TATTTCCGCTTGAATACCTTAGCCTCTAGCTTATGTTGTTAGTAGATAGACACAGG 2280  
DB |||||  
QY 2281 GATCTCTGTCACCGAGGGCTTAAATTTGAAGGGCTTTAAATATATAAATGGGTCAAATA 2340  
DB GATCTCTGTCACCGAGGGCTTAAATTTGAAGGGCTTTAAATATATAAATGGGTCAAATA 2340  
QY 2341 GAAATTCACCATCAGGTGGAAGGACCTCAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400  
DB GAAATTCACCATCAGGTGGAAGGACCTCAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400  
QY 2401 TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAATTA 2460  
DB TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAATTA 2460  
QY 2461 ATGACAGGTTCAGATATCTTCAGGGTGTCTATTGCTGGGGCCCGACGTCTCTGTGATC 2520  
DB ATGACAGGTTCAGATATCTTCAGGGTGTCTATTGCTGGGGCCCGACGTCTCTGTGATC 2520  
QY 2521 TTCTATGATACAGGATACAGGAACTTATATGAGGTTCACCTGACCAAGATGAACAGGC 2580  
DB TTCTATGATACAGGATACAGGAACTTATATGAGGTTCACCTGACCAAGATGAACAGGC 2580  
QY 2581 TATTACTTAGATCTGTGGCCATGCAAGCAGAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA 2640  
DB TATTACTTAGATCTGTGGCCATGCAAGCAGAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA 2640  
QY 2641 CTGCTCTTACATGTTTCTGGAAGAGAAATGTCATTTTGCACATACAGATATATTAATG 2700  
DB CTGCTCTTACATGTTTCTGGAAGAGAAATGTCATTTTGCACATACAGATATATTAATG 2700  
QY 2701 AGTTTTTTAGTAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGAGACAC 2760  
DB AGTTTTTTAGTAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGAGACAC 2760  
QY 2761 AGCAATAGAGTTCTGTAATCGGAGAACATTTATGAACTGCACTTTTGCATCTTCAA 2820  
DB AGCAATAGAGTTCTGTAATCGGAGAACATTTATGAACTGCACTTTTGCATCTTCAA 2820  
QY 2821 GAAAACCTTGATCACGTATGCTGCTCTAAAGTGATATATTTTGAACCTGTGTAGAAC 2880  
DB GAAAACCTTGATCACGTATGCTGCTCTAAAGTGATATATTTTGAACCTGTGTAGAAC 2880  
QY 2881 TCTCTGTATACACTGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAG 2940  
DB TCTCTGTATACACTGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAG 2940  
QY 2941 ATGATCATCATATTTTATGATACCTGCGATGATCACTACTCTGAAATATAATGTGGTG 3000  
DB ATGATCATCATATTTTATGATACCTGCGATGATCACTACTCTGAAATATAATGTGGTG 3000  
QY 3001 CATTGACAGGGTCTACGGTTTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAA 3060  
DB CATTGACAGGGTCTACGGTTTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAA 3060  
QY 3061 TCAATATGATATATTTCTGAGAGACCCAGCATACATAAGATTAATAAAAAAAA 3120  
DB TCAATATGATATATTTCTGAGAGACCCAGCATACATAAGATTAATAAAAAAAA 3120

RESULT 3  
ABK12892  
ID ABK12892 standard; cDNA; 3106 BP.  
XX  
AC ABK12892;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Human protease PRTS-9 cDNA sequence.  
XX  
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;

KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
KW cell proliferative disorder; developmental disorder; epilepsy;  
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
KW reproductive disorder; endometriosis; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 203..2851  
FT /\*tag= a  
FT /product= "Human protease PRTS-9"  
XX  
PN WO200198468-A2.  
XX  
PD 27-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019178.  
XX  
PR 16-JUN-2000; 2000US-0212336P.  
PR 22-JUN-2000; 2000US-0213955P.  
PR 29-JUN-2000; 2000US-0215396P.  
PR 07-JUL-2000; 2000US-0216821P.  
PR 14-JUL-2000; 2000US-0218946P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;  
PI Delegeane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;  
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
PI Kallick DA;  
XX  
DR WPI; 2002-090437/12.  
XX P-PSDB; AAU74749.  
XX  
PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in  
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative  
PT (e.g. cancer) disorders.  
XX  
PS Claim 5; Page 166-167; 177pp; English;  
XX  
CC The present invention relates to twenty one new human proteases, referred  
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the  
CC invention are useful in the diagnosis, treatment and prevention of  
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's  
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial  
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency  
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.  
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,  
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's  
CC disease and reproductive e.g. infertility and endometriosis disorders.  
CC Numerous other examples of each disorder are given in the specification.  
CC The present nucleic acid sequence encodes the human protease PRTS-9  
XX protein of the invention  
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;

Query Match 96.2%; Score 3003; DB 6; Length 3106;  
Best Local Similarity 99.9%; Pred. No; 0;  
Matches 3103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 13 CTCGAGGCCAAGGCCGCTCTACTGCGCGCGCTGCTTCTTAGTCCCGCTTCGCGCCT 72  
DB |||||  
DB 2 CTCGAGGCCAAGGCCGCTCTACTGCGCGCGCTTCTTAGTCCCGCTTCGCGCCT 61  
QY 73 GGGTTGTACCG 132  
DB |||||  
DB 62 GGGTTGTACCG 121  
QY 133 CGCAGCATGAGCG 192  
DB |||||  
DB 122 CGCAGCATGAGCG 181

QY 193 CCGGGGGAAGAAAATGCAACATGGCAGCAGCAATGGAAACAGAACACAGCTCGGTGTTGAG 252  
DB 182 CCGGGGGAAGAAAATGCAACATGGCAGCAGCAATGGAAACAGAACACAGCTGGGTGTTGAG 241  
QY 253 ATATTTGAAACTGCGGACTGTGAGGAGATATTGAATCACAGATCGGCTAAATTTGAG 312  
DB 242 ATATTTGAAACTGCGGACTGTGAGGAGATATTGAATCACAGATCGGCTAAATTTGAG 301  
QY 313 CCTTTTATGTTGAGCGGTATCTCGAGTCAAGCTTTAAAGACTGCTTCCCGATACAGA 372  
DB 302 CCTTTTATGTTGAGCGGTATCTCGAGTCAAGCTTTAAAGACTGCTTCCCGATACAGA 361  
QY 373 AAATATCATGGCTACATGATGGCTAAGGCACCACTATGATTTTCAATGTTTGTGAAAGGAAT 432  
DB 362 AAATATCATGGCTACATGATGGCTAAGGCACCACTATGATTTTCAATGTTTGTGAAAGGAAT 421  
QY 433 GATTCAGATGGAACCTCATTCAGACAGAACTCTATTAACCTTGGCAGTCTGTGGTGAACAGA 492  
DB 422 GATTCAGATGGAACCTCATTCAGACAGAACTCTATTAACCTTGGCAGTCTGTGGTGAACAGA 481  
QY 493 GAAATACACTGTTTATCTGAAATTCCTCAAAACTATCAATAGACAGCAGTCTTAATG 552  
DB 482 GAAATACACTGTTTATCTGAAATTCCTCAAAACTATCAATAGACAGCAGTCTTAATG 541  
QY 553 CTCTCTGGAGGCTCTTTTGGATCTTTTTCAGGCACACTGGACTATGGAATGATATCT 612  
DB 542 CTCTCTGGAGGCTCTTTTGGATCTTTTTCAGGCACACTGGACTATGGAATGATATCT 601  
QY 613 CGAGAAGAACTATTAAGAGAAAGAAACGATTTGGAACAGTCGGAATTTGCTTCTTAC 672  
DB 602 CGAGAAGAACTATTAAGAGAAAGAAACGATTTGGAACAGTCGGAATTTGCTTCTTAC 661  
QY 673 GATTATCACAAGAGTGGACATTTCTGTTTTCAGCGGTAGTGGAAATTTATCAGGTA 732  
DB 662 GATTATCACAAGAGTGGACATTTCTGTTTTCAGCGGTAGTGGAAATTTATCAGGTA 721  
QY 733 AAAGATCGAGGCGCAAGAGATTAAGCAACCTTTAAGCCCAATCTAGTGGAACT 792  
DB 722 AAAGATCGAGGCGCAAGAGATTAAGCAACCTTTAAGCCCAATCTAGTGGAACT 781  
QY 793 AGTTGTCGCAACATCGATGATCCAAATTTATGCCCGCTGATCCAGACTGGATGCT 852  
DB 782 AGTTGTCGCAACATCGATGATCCAAATTTATGCCCGCTGATCCAGACTGGATGCT 841  
QY 853 TTTATACATAGCAACGATTTTGGATATCTAACATCGTAAACAGAGAAAGAGACTC 912  
DB 842 TTTATACATAGCAACGATTTTGGATATCTAACATCGTAAACAGAGAAAGAGACTC 901  
QY 913 ACTATGTGCACATGAGCTAGCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCT 972  
DB 902 ACTATGTGCACATGAGCTAGCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCT 961  
QY 973 ACCTTTGTTCTCAAGAAAGATTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCT 1032  
DB 962 ACCTTTGTTCTCAAGAAAGATTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCT 1021  
QY 1033 GAAACAACTCCCAAGTGGTGGTAAAATCTTAGAATTTCTATAGAAATATGATGATCT 1092  
DB 1022 GAAACAACTCCCAAGTGGTGGTAAAATCTTAGAATTTCTATAGAAATATGATGATCT 1081  
QY 1093 GAGGTGGAATTTATCATGTTACATCCCTATGTTGGAACAGAGGCGGAGATTCATTC 1152  
DB 1082 GAGGTGGAATTTATCATGTTACATCCCTATGTTGGAACAGAGGCGGAGATTCATTC 1141  
QY 1153 CGTTATCCCTAAACAGGTACAGCAATCTTAAGTCACTTTTAAGTATGTCAGAAATATG 1212  
DB 1142 CGTTATCCCTAAACAGGTACAGCAATCTTAAGTCACTTTTAAGTATGTCAGAAATATG 1201  
QY 1213 ATGTGCTGAAGGAGATCATAGATGTCATAGATGAAGCACTAAATCAACCTTTTGG 1272  
DB 1202 ATGTGCTGAAGGAGATCATAGATGTCATAGATGAAGCACTAAATCAACCTTTTGG 1261

QY 1273 ATTCTATTGAAAGGATTGAATATATTGCCAGAGCTGGATGACTCCTGAGGGAATAATAT 1332  
DB 1262 ATTCTATTGAAAGGATTGAATATATTGCCAGAGCTGGATGACTCCTGAGGGAATAATAT 1321  
QY 1333 GCTTGGTCCATCTCTACATGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCT 1392  
DB 1322 GCTTGGTCCATCTCTACATGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCT 1381  
QY 1393 GAAATATTTATCCCAAGTGAAGATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTG 1452  
DB 1382 GAAATATTTATCCCAAGTGAAGATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTG 1441  
QY 1453 CCTGATTTCTGTGACCGCACTAATTTATGTAAGAAACAAACAGACATCTGGATAAATATC 1512  
DB 1442 CCTGATTTCTGTGACCGCACTAATTTATGTAAGAAACAAACAGACATCTGGATAAATATC 1501  
QY 1513 CATGACATCTTTTCAATGTTTTTCCCAAGTCAAGAGAGGAAATTTGAGTTTTTTTTGGCC 1572  
DB 1502 CATGACATCTTTTCAATGTTTTTCCCAAGTCAAGAGAGGAAATTTGAGTTTTTTTTGGCC 1561  
QY 1573 TCTGAATGCAAAACAGGTTTTCCGTCAATTTATACAAATTTACATCTATTTTAAAGGAAGC 1632  
DB 1562 TCTGAATGCAAAACAGGTTTTCCGTCAATTTATACAAATTTACATCTATTTTAAAGGAAGC 1621  
QY 1633 AAATATAAACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAA 1692  
DB 1622 AAATATAAACGATCCAGTGGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAA 1681  
QY 1693 GAGGAGTAGCAATACCACTGGTGGTGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATC 1752  
DB 1682 GAGGAGTAGCAATACCACTGGTGGTGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATC 1741  
QY 1753 CAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGCACCAAGACTCCCTTTTAGAG 1812  
DB 1742 CAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGCACCAAGACTCCCTTTTAGAG 1801  
QY 1813 CATCACTGTGACGTAGTCAAGTTACGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGT 1872  
DB 1802 CATCACTGTGACGTAGTCAAGTTACGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGT 1861  
QY 1873 GGCTACTCACTTTCTGTGTCATCAGTCAAGTGTGACTTCTTTTATAAGTAAAGTATAGT 1932  
DB 1862 GGCTACTCACTTTCTGTGTCATCAGTCAAGTGTGACTTCTTTTATAAGTAAAGTATAGT 1921  
QY 1933 AACCAAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCA 1992  
DB 1922 AACCAAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCA 1981  
QY 1993 ACTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGAGTCTCTTCTTGAC 2052  
DB 1982 ACTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGAGTCTCTTCTTGAC 2041  
QY 2053 TATACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATG 2112  
DB 2042 TATACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATG 2101  
QY 2113 CTCTACAGGCTCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTTCATATAT 2172  
DB 2102 CTCTACAGGCTCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTTCATATAT 2161  
QY 2173 GGTGCTCTCAGGTCAGTGGTGAATAAATCGGTTTTTAAAGGAGTCAAGTATTTCCGCTTG 2232  
DB 2162 GGTGCTCTCAGGTCAGTGGTGAATAAATCGGTTTTTAAAGGAGTCAAGTATTTCCGCTTG 2221  
QY 2233 AATACCTAGGCTCTCTAGGTTATGTGGTTGTAGTAGACAAACAGGGGATCCTGTAC 2292  
DB 2222 AATACCTAGGCTCTCTAGGTTATGTGGTTGTAGTAGACAAACAGGGGATCCTGTAC 2281  
QY 2293 CGAGGCTTAAATTTGAAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGAT 2352  
DB 2282 CGAGGCTTAAATTTGAAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGAT 2341  
QY 2353 CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTAGATCGTGTG 2412

Db 2342 CAGGTGAGAGCTCAATATCTAGCTTCGATATGATTTCAATGACTTAGATCGGTG 2401  
QY 2413 GGATCCACGGCTGCTCTATGAGGATACCTCCCTGATGCGATTAATGACAGGTCA 2472  
Db 2402 GGATCCACGGCTGCTCTATGAGGATACCTCCCTGATGCGATTAATGACAGGTCA 2461  
QY 2473 GATATCTTCAGGGTGTCTATGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA 2532  
Db 2462 GATATCTTCAGGGTGTCTATGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA 2521  
QY 2533 GGATACAGGAACGTTATATGAGGTCACTCTGACCAAGATGAACAGGGCTATTTAGGA 2592  
Db 2522 GGATACAGGAACGTTATATGAGGTCACTCTGACCAAGATGAACAGGGCTATTTAGGA 2581  
QY 2593 TCTGTGGCATGCAAGCAAGATGTTCCCTCTGAACCAAAATCGTTTACTGCTTACAT 2652  
Db 2582 TCTGTGGCATGCAAGCAAGATGTTCCCTCTGAACCAAAATCGTTTACTGCTTACAT 2641  
QY 2653 GGTTCCTCGATGAGAAATGTCATTTTGACATACAGTATATTAATCTAGTCTTTTAGTG 2712  
Db 2642 GGTTCCTCGATGAGAAATGTCATTTTGACATACAGTATATTAATCTAGTCTTTTAGTG 2701  
QY 2713 AGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTT 2772  
Db 2702 AGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGAGTT 2761  
QY 2773 CCTCAATCGGGAGCAATATGATGATCTTTTGCACTACCTTTCAAGAAAACCTTGGGA 2832  
Db 2762 CCTCAATCGGGAGCAATATGATGATCTTTTGCACTACCTTTCAAGAAAACCTTGGGA 2821  
QY 2833 TCAGTATGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATAC 2892  
Db 2822 TCAGTATGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATAC 2881  
QY 2893 ACTGGCTATTAAACCAATGAGAGGTTTAATCAACAGAAAACAGAAATGATCATCAC 2952  
Db 2882 ACTGGCTATTAAACCAATGAGAGGTTTAATCAACAGAAAACAGAAATGATCATCAC 2941  
QY 2953 ATTTTGATCTGCTGCTAACTACTCTCTGAAATAATGTTGCTCCATGAGGGGT 3012  
Db 2942 ATTTTGATCTGCTGCTAACTACTCTCTGAAATAATGTTGCTCCATGAGGGGT 3001  
QY 3013 CTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCAATGCTCAAAATCAAAATGATACA 3072  
Db 3002 CTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCAATGCTCAAAATCAAAATGATACA 3061  
QY 3073 TATTCCTGAGAGCCCAAGCAATACCAATGATTAATAAATAA 3117  
Db 3062 TATTCCTGAGAGCCCAAGCAATACCAATGATTAATAAATAA 3106

RESULT 4  
AAH99934  
ID AAH99934 standard; cDNA; 3143 BP.

AC AAH99934;

DT 12-APR-2002 (first entry)

DE cDNA encoding 21953 human prollyl oligopeptidase.

KW 21953 prollyl oligopeptidase; human; proline; endopeptidase; cancer;  
KW cardiovascular disease; autoimmune disease; atopic allergy;  
KW neuronal disorder; vascular disease; prostate disorder; cytostatic;  
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;  
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;  
KW Grave's disease; neuronal disorder; demyelinating disease; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH 229. .2877  
FT CDS

/tag= a  
/product= "21953 prollyl oligopeptidase"  
/note= "This region is specifically claimed in claim 2"  
WO200179473-A2.  
XX 25-OCT-2001.  
XX 11-APR-2001; 2001WO-US040483.  
XX 18-APR-2000; 2000US-0197508P.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Meyers RA, Williamson M;  
XX WPI; 2002-034353/04.  
XX P-PSDB; AAG78415.  
XX New polypeptides 21953, member of human prollyl oligopeptidase family,  
XX useful as diagnostic targets and therapeutic agents for controlling  
XX cancer, lymphoma and leukemia.  
XX Claim 7; Page 100-102; 121pp; English.  
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.  
XX Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,  
XX antithyroid, dermatological, antipsoriatic, antiasthmatic,  
XX ophthalmological, antiinflammatory, nootropic, antianginal, cardiac,  
XX anticonvulsant, gynaecological, vasotropic, antianginal, cardiac,  
XX antiatherosclerotic, anorectic and metabolic in its action. Uses include  
XX gene therapy, expression or activity of 21953 protein modulator, it is  
XX useful for identifying a compound which binds to it and can be used in  
XX preventing, treating or detecting a cellular proliferative or  
XX differentiative disorder. The 21953 molecules can act as novel diagnostic  
XX targets and therapeutic agents for controlling disorders associated with  
XX the aberrant activity or degradation of peptide hormones e.g., disorders  
XX associated with cell differentiation and proliferation such as cancer,  
XX immune function, reproductive, neurological and cardiovascular function.  
XX The 21953 molecules are thus useful for treating and preventing cellular  
XX proliferative and differentiative disorders, haematopoietic neoplastic  
XX disorders, immune disorders such as autoimmune diseases, diabetes  
XX mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,  
XX neuronal disorders, demyelinating diseases, vascular disorders and  
XX metabolism or pain disorders. This sequence represents the cDNA encoding  
XX sequence of 21953 human prollyl oligopeptidase

SQ Sequence 3143 BP; 943 A; 644 C; 712 G; 844 T; 0 U; 0 Other;

Query Match 94.8%; Score 2957; DB 6; Length 3143;  
Best Local Similarity 99.9%; Pred. No; 0;  
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 GCCTCCGAGCCCAAGCCGCTGCTACTGACCGCGCTGCTTCTTAGTGCGCGGTTCCGCCG 70  
Db 26 GCCTCCGAGCCCAAGCCGCTGCTACTGACCGCGCTGCTTCTTAGTGCGCGGTTCCGCCG 85  
QY 71 CTGGGTTGTCACTCCGCGCCGCCCGCCGAGGAGCACTCAACACGAGACCGGAGTGAGGC 130  
Db 86 CTGGGTTGTCACTCCGCGCCGCCCGCCGAGGAGCACTCAACACGAGACCGGAGTGAGGC 145  
QY 131 GCGCAGCATGACCGCGCCGCGCGCTCCATAGCGGACGTCGCGGCGGTCGCGCGGG 190  
Db 146 GCGCAGCATGACCGCGCCGCGCGCTCCATAGCGGACGTCGCGGACGTCGCGGCGG 205  
QY 191 GCCCGGGGAGGAAATGCAACATGCGAGCAGCAATGGAACAGAAACAGCTGGGTGTG 250  
Db 206 GCCCGGGGAGGAAATGCAACATGCGAGCAGCAATGGAACAGAAACAGCTGGGTGTG 265  
QY 251 AGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAAGATCGGCCCTAAATGG 310  
Db 266 AGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAAGATCGGCCCTAAATGG 325

QY 311 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGCCGATACCA 370  
Db 326 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGCCGATACCA 385  
QY 371 GAAATATCATGCGCTACATGATGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGA 430  
Db 386 GAAATATCATGCGCTACATGATGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGA 445  
QY 431 ATGATCCAGATGGACCTCATTTCAGACAGAAATCTATTACCTTCCCATGTCCTGGTGAACCA 490  
Db 446 ATGATCCAGATGGACCTCATTTCAGACAGAAATCTATTACCTTCCCATGTCCTGGTGAACCA 505  
QY 491 GAGAAATACACTGCTTTTATTCCTGAATTTCCCAAACTATCAATAGACGACGCTCTTAA 550  
Db 506 GAGAAATACACTGCTTTTATTCCTGAATTTCCCAAACTATCAATAGACGACGCTCTTAA 565  
QY 551 TGCTCTCTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGACATATGGAATGATT 610  
Db 566 TGCTCTCTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGACATATGGAATGATT 625  
QY 611 CTCGAGAGAGAACTATTAAGAGAAAGAAACGCAATGGAAACAGTCGGAATTCCTCTT 670  
Db 626 CTCGAGAGAGAACTATTAAGAGAAAGAAACGCAATGGAAACAGTCGGAATTCCTCTT 685  
QY 671 ACGATATACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACG 730  
Db 686 ACGATATACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACG 745  
QY 731 TAAAGATGGAGGCGCCAAAGGATTTACGAAACAACTTTAAGGCCCAATCTAGTGGAAA 790  
Db 746 TAAAGATGGAGGCGCCAAAGGATTTACGAAACAACTTTAAGGCCCAATCTAGTGGAAA 805  
QY 791 CTAGTTGTCCTCAACATACGAGTGGATCCAAATTTATGCCCCGCTGATCCAGACTGGATTG 850  
Db 806 CTAGTTGTCCTCAACATACGAGTGGATCCAAATTTATGCCCCGCTGATCCAGACTGGATTG 865  
QY 851 CTTTATATACATAGCAACGATATTTGGATATCTAAATCTGTAACATGTAACAGAGAAAGAGGAC 910  
Db 866 CTTTATATACATAGCAACGATATTTGGATATCTAAATCTGTAACATGTAACAGAGAAAGAGGAC 925  
QY 911 TCACCTATGTGCAATAGCTAGCCAAACATGGAAAGAGATGCCAGACTCAGCTGGAGTCG 970  
Db 926 TCACCTATGTGCAATAGCTAGCCAAACATGGAAAGAGATGCCAGACTCAGCTGGAGTCG 985  
QY 971 CTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTCCTGGCTATTTGGTGGTCCAAAG 1030  
Db 986 CTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTCCTGGCTATTTGGTGGTCCAAAG 1045  
QY 1031 CTGAACCACTCCAGTGGTGTAAATCTTAGAATCTATATGAGAAATGATGAAT 1090  
Db 1046 CTGAACCACTCCAGTGGTGTAAATCTTAGAATCTATATGAGAAATGATGAAT 1105  
QY 1091 CTGAGTGGAAATTTATGTTATCATCCCTATGTTGGAAACAGAGGCGCAGATTCAAT 1150  
Db 1106 CTGAGTGGAAATTTATGTTATCATCCCTATGTTGGAAACAGAGGCGCAGATTCAAT 1165  
QY 1151 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAA 1210  
Db 1166 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAA 1225  
QY 1211 TGATTGATGCTGAAGAGGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTG 1270  
Db 1226 TGATTGATGCTGAAGAGGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTG 1285  
QY 1271 AGATTCTATTGAAGAGGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGGAAAT 1330  
Db 1286 AGATTCTATTGAAGAGGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGGAAAT 1345  
QY 1331 ATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCAC 1390  
Db 1346 ATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCAC 1405  
QY 1391 CTGAATATTATTTATCCAGTAGAAGATGATGTTATGGAAAGGCGAGAGACTCATTGAGTCAG 1450

Db 1406 CTGAATATTATTTATCCAGTAGAAGATGATTTATGAAAGGCGAGAGCTCATTGAGTCAG 1465  
QY 1451 TGCCCTGATCTCTGAGCGCGCACTAATTTATCTATGAAAGAAACAACAGACATCTGGATAAATA 1510  
Db 1466 TGCCCTGATCTCTGAGCGCGCACTAATTTATCTATGAAAGAAACAACAGACATCTGGATAAATA 1525  
QY 1511 TCCATGACATCTTTTCATGTTTTCCTCCAAAGTCAAGAGAGAAATTTGAGTTTATTTTTG 1570  
Db 1526 TCCATGACATCTTTTCATGTTTTCCTCCAAAGTCAAGAGAGAAATTTGAGTTTATTTTTG 1585  
QY 1571 CCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTAACATCTATTTTAAAGGAAA 1630  
Db 1586 CCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTAACATCTATTTTAAAGGAAA 1645  
QY 1631 GCAAAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTCCAAGTGTCTCTATCA 1690  
Db 1646 GCAAAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTCCAAGTGTCTCTATCA 1705  
QY 1691 AAGAGGAGATAGCAATTTAACCAGTGGTGAATGGGAAAGTTCTTGGCCGGCATGGATCTAATA 1750  
Db 1706 AAGAGGAGATAGCAATTTAACCAGTGGTGAATGGGAAAGTTCTTGGCCGGCATGGATCTAATA 1765  
QY 1751 TCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTTTTAG 1810  
Db 1766 TCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTTTTAG 1825  
QY 1811 AGCATCACCTGTACGTAGTACGTTAGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACC 1870  
Db 1826 AGCATCACCTGTACGTAGTACGTTAGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACC 1885  
QY 1871 GTGGCTACTCACATCTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATA 1930  
Db 1886 GTGGCTACTCACATCTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATA 1945  
QY 1931 GTAACCAAGAGATCCACACTCTGTGTCCCTTTTACAGCTATCAAGTCTCTGGAAGTACC 1990  
Db 1946 GTAACCAAGAGATCCACACTCTGTGTCCCTTTTACAGCTATCAAGTCTCTGGAAGTACC 2005  
QY 1991 CAACCTTGCAAAACAAGGAATTTTGGGCCACATTTTGGATTCAGAGAGTCTCTTCCCTG 2050  
Db 2006 CAACCTTGCAAAACAAGGAATTTTGGGCCACATTTTGGATTCAGAGAGTCTCTTCCCTG 2065  
QY 2051 ACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGATGGGA 2110  
Db 2066 ACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGATGGGA 2125  
QY 2111 TGCTCTACAGGCTCATGATCTACAGCTGGAAGAAATATCTCTACGTGCTGTTCATAT 2170  
Db 2126 TGCTCTACAGGCTCATGATCTACAGCTGGAAGAAATATCTCTACGTGCTGTTCATAT 2185  
QY 2171 ATGGTGGTCTCAGGTGCAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2230  
Db 2186 ATGGTGGTCTCAGGTGCAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2245  
QY 2231 TGAATACCCCTAGCCCTCTCTAGGTTATGTTGGTGTAGTAGAGACAACAGGGGATCTGTG 2290  
Db 2246 TGAATACCCCTAGCCCTCTCTAGGTTATGTTGGTGTAGTAGAGACAACAGGGGATCTGTG 2305  
QY 2291 ACCGAGGGCTTAAATTTTGAAGCGCTTTAAATATAAAATGGGTCAAATAGAAATTTGACG 2350  
Db 2306 ACCGAGGGCTTAAATTTTGAAGCGCTTTAAATATAAAATGGGTCAAATAGAAATTTGACG 2365  
QY 2351 ATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTAGATCGTG 2410  
Db 2366 ATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTAGATCGTG 2425  
QY 2411 TGGGGATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGGATTAATGACAGGT 2470  
Db 2426 TGGGGATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGGATTAATGACAGGT 2485  
QY 2471 CAGATATCTTCAGGGTGTCTATTCGCGGCCCGAGTCACTCTGTGGATCTCTATGATA 2530

Db 2486 CAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATA 2545  
Qy 2531 CAGGATACCGGAACGTTATATGAGTCACTGACCGAGATGAACAGGGCTATTAATTAG 2590  
Db 2546 CAGGATACCGGAACGTTATATGAGTCACTGACCGAGATGAACAGGGCTATTAATTAG 2605  
Qy 2591 GATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTCTCTTAC 2650  
Db 2606 GATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTCTCTTAC 2665  
Qy 2651 ATGTTTCTCGGATGAGATGTCATTTTGACATACCAAGTATTAATTACTGAGTTTTTAG 2710  
Db 2666 ATGTTTCTCGGATGAGATGTCATTTTGACATACCAAGTATTAATTACTGAGTTTTTAG 2725  
Qy 2711 TGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAGCATAGAG 2770  
Db 2726 TGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAGCATAGAG 2785  
Qy 2771 TTCTGTAATCGGAGAACATTATGAATGCTGATCTTTTGCATCTTCAAGAAACCTTG 2830  
Db 2786 TTCTGTAATCGGAGAACATTATGAATGCTGATCTTTTGCATCTTCAAGAAACCTTG 2845  
Qy 2831 GATCAGTATGCTCTCTAAAGTATATTTTGACCTGTCTAGACTCTCTGAT 2890  
Db 2846 GATCAGTATGCTCTCTAAAGTATATTTTGACCTGTCTAGACTCTCTGAT 2905  
Qy 2891 ACATGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATC 2950  
Db 2906 ACATGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATC 2965  
Qy 2951 ACATTTGTACCTGCGATGTAACATCTACTCTGAAATAAATGTTGTCGTCAGG 3010  
Db 2966 ACATTTGTACCTGCGATGTAACATCTACTCTGAAATAAATGTTGTCGTCAGG 3025  
Qy 3011 GTCTACCGTTTGTGTAGTAATCTAATACCTTTAACCACATGCTCAAAATCAATGATA 3070  
Db 3026 GTCTACCGTTTGTGTAGTAATCTAATACCTTTAACCACATGCTCAAAATCAATGATA 3085  
Qy 3071 CATATCTGAGAGACCCAGCAATACATAAGAAATTAATAAAAAA 3120  
Db 3086 CATATCTGAGAGACCCAGCAATACATAAGAAATTAATAAAAAA 3135

RESULT 5

ID ABK83327 standard; cDNA; 4829 BP.

XX AC ABK83327;

XX DT 12-AUG-2002 (first entry)

XX DE cDNA encoding human DPRP-1 splice variant #3.

XX KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;  
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;  
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;  
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;  
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;  
KW dyskinnesia; reproductive disorder; inflammatory disorder;  
KW metabolic disorder; gene; ss.

XX OS Homo sapiens.

XX PN WO200231134-A2.

XX XX 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US031874.

XX PR 12-OCT-2000; 2000US-0240117P.

XX PA (FERR ) FERRING BV.

PI Qi S, Akinsanya KO, Riviere PJ, Junien J;  
XX WPI: 2002-444178/47.  
DR P-PSDB; ABG61596.  
XX  
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding  
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and  
PT viral infections, cancers, allergies, neurological disorders, or pain.  
XX  
PS Disclosure; Page 65-66; 113pp; English.  
XX  
CC The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPPiV)-related proteins  
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic  
CC acids encoding them are useful for treating infections such as fungal,  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322  
CC -ABK83343 encode human DPRP proteins  
XX

SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;

Query Match 92.6%; Score 2889; DB 6; Length 4829;  
Best Local Similarity 99.9%; Pred. No. 0;

Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 1 AAGTGTCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTCCCGCGCTGTTTATGTCGG 60  
Db 1 AAGTGTCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTCCCGCGCTGTTTATGTCGG 60  
Qy 61 CGTTTCGGCGCTGGTGTGTACCGCGCGCCGCGAGAGCCACTGCAACCGAGGACCG 120  
Db 61 CGTTTCGGCGCTGGTGTGTACCGCGCGCCGCGAGAGCCACTGTCACCCAGGACCG 120  
Qy 121 GAGTGGAGGCGCGCAGCATGAAGCGCGCCGCTCCATAGCGCACCTCCGGACGG 180  
Db 121 GAGTGGAGGCGCGCAGCATGAAGCGCGCCGCTCCATAGCGCACCTCCGGACGG 180  
Qy 181 TCCGGGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG 240  
Db 181 TCCGGGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGAACAG 240  
Qy 241 CTGGGTCTTGAGATATTTGAAACTGCGGA CTGTGAGGAGAAATATTGAATCAGAGATCGG 300  
Db 241 CTGGGTCTTGAGATATTTGAAACTGCGGA CTGTGAGGAGAAATATTGAATCAGAGATCGG 300  
Qy 301 CCTAAATTGGAGCCTTTTATGTTGAGCGGTATCTGGAGTCAGCTTAAAGCTGCTT 360  
Db 301 CCTAAATTGGAGCCTTTTATGTTGAGCGGTATCTGGAGTCAGCTTAAAGCTGCTT 360  
Qy 361 GCCGATACCGAAAAATATCATGCTACATGCTTAAAGCCACCATGATTTTCATGTTT 420  
Db 361 GCCGATACCGAAAAATATCATGCTACATGCTTAAAGCCACCATGATTTTCATGTTT 420  
Qy 421 GTGAGGAGGATGATCCAGATGGACCTCTTTCAGACAGAAATCTATTACCTTGGCCATGCT 480  
Db 421 GTGAGGAGGATGATCCAGATGGACCTCTTTCAGACAGAAATCTATTACCTTGGCCATGCT 480  
Qy 481 GTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTTCCCAAAATATCAATAGACA 540  
Db 481 GTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTTCCCAAAATATCAATAGACA 540  
Qy 541 GCAGTCTTAATGCTCTCTTGGAGACCTCTTTTGGATCTTTTTCAGGCAACACTGCACTAT 600  
Db 541 GCAGTCTTAATGCTCTCTTGGAGACCTCTTTTGGATCTTTTTCAGGCAACACTGCACTAT 600  
Qy 601 GGAATGTATTCTCGAGAGAGAAAGAACTATTAGAGAGAAAGAAACCGCATTCGACACATCGGA 660



Db 601 GGAATGTATTCTCAGAGAGAGACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCGGG 660  
Qy 661 ATTGCTTCTTACGATTATACCCAGGAGTGGAGCAATTTCTGTTTCAAGCCGGTAGTGGG 720  
Db 661 ATTGCTTCTTACGATTATACCCAGGAGTGGAGCAATTTCTGTTTCAAGCCGGTAGTGGG 720  
Qy 721 ATTATACAGCTTAAAGATGAGAGGCGCCACAGGATTTACGCAACAACCTTTTAAAGCCCAAT 780  
Db 721 ATTATACAGCTTAAAGATGAGAGGCGCCACAGGATTTACGCAACAACCTTTTAAAGCCCAAT 780  
Qy 781 CTAGTGGAACTAGTGTGTCACACATACGGATGGATTCGCAAAATTAATGCCCCGCTGATCCA 840  
Db 781 CTAGTGGAACTAGTGTGTCACACATACGGATGGATTCGCAAAATTAATGCCCCGCTGATCCA 840  
Qy 841 GACTGGATTTGCTTTTATACATAGCAACAGATATTTGGATATCTAAACATCGTAAACAGAGAA 900  
Db 841 GACTGGATTTGCTTTTATACATAGCAACAGATATTTGGATATCTAAACATCGTAAACAGAGAA 900  
Qy 901 GAAAGGAGACTCACTTATGTGCAAAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCA 960  
Db 901 GAAAGGAGACTCACTTATGTGCAAAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCA 960  
Qy 961 GCTGGAGTCTGCTACCTTTGTTCTCCAGAGAGAAATTTGATAGATATCTGGCTATTGGTGG 1020  
Db 961 GCTGGAGTCTGCTACCTTTGTTCTCCAGAGAGAAATTTGATAGATATCTGGCTATTGGTGG 1020  
Qy 1021 TGTCGAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAAAGATTTCTATATGAGAA 1080  
Db 1021 TGTCGAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAAAGATTTCTATATGAGAA 1080  
Qy 1081 AATGATCAATCTCAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAACAGAGG 1140  
Db 1081 AATGATCAATCTCAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAACAGAGG 1140  
Qy 1141 GCAGATTCATTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG 1200  
Db 1141 GCAGATTCATTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG 1200  
Qy 1201 TCAGAAATATGATGCTGCAAGAGGATCATAGATGTCATAGATAAGAACTAAAT 1260  
Db 1201 TCAGAAATATGATGCTGCAAGAGGATCATAGATGTCATAGATAAGAACTAAAT 1260  
Qy 1261 CAACCTTTTGAGATTTCTATTTGAAGAGTGAATATATGTCAGAGCTGGATGCTCT 1320  
Db 1261 CAACCTTTTGAGATTTCTATTTGAAGAGTGAATATATATGTCAGAGCTGGATGCTCT 1320  
Qy 1321 GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380  
Db 1321 GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380  
Qy 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC 1440  
Db 1381 TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC 1440  
Qy 1441 ATTGAGTCAGTGCCTGATTCTGTGACGCCACTAATTTATCTATGAAGAAACAGACATC 1500  
Db 1441 ATTGAGTCAGTGCCTGATTCTGTGACGCCACTAATTTATCTATGAAGAAACAGACATC 1500  
Qy 1501 TGGATAAATATCATGACATCTTTTCATGTTTTTCCCAAAAGTCCAGAGGAGAAATGAG 1560  
Db 1501 TGGATAAATATCATGACATCTTTTCATGTTTTTCCCAAAAGTCCAGAGGAGAAATGAG 1560  
Qy 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCAATTTATACAAATTTATCTATT 1620  
Db 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCAATTTATACAAATTTATCTATT 1620  
Qy 1621 TTAAGGAAAGCAATATTAACGATCCAGTGGTGGCTGCTCCAGTGAATTTCAAG 1680  
Db 1621 TTAAGGAAAGCAATATTAACGATCCAGTGGTGGCTGCTCCAGTGAATTTCAAG 1680  
Qy 1681 TGTCTATCAAGAGGATAGCAATTTACCAGTGGTGAATGGGAATTTCTTGGCCGGCAT 1740  
Db 1681 TGTCTATCAAGAGGATAGCAATTTACCAGTGGTGAATGGGAATTTCTTGGCCGGCAT 1740

Db 1681 TGTCTATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAATTTCTTGGCCGGCAT 1740  
Qy 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGACCAAAAGAC 1800  
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGACCAAAAGAC 1800  
Qy 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGG 1860  
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGG 1860  
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCAGCTGACTTCTTTATA 1920  
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCAGCTGACTTCTTTATA 1920  
Qy 1921 AGTAAGTATAGTAACAGAGAAATCCACATCTGTGTGCCCTTACAGCTATCAAGTCCT 1980  
Db 1921 AGTAAGTATAGTAACAGAGAAATCCACATCTGTGTGCCCTTACAGCTATCAAGTCCT 1980  
Qy 1981 GAAGATGACCCAACTTGCAGAAACAGAGAAATTTGGGCCACATTTTGGATTCAGCAGGT 2040  
Db 1981 GAAGATGACCCAACTTGCAGAAACAGAGAAATTTGGGCCACATTTTGGATTCAGCAGGT 2040  
Qy 2041 CCTCTTCCGCTGATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA 2100  
Db 2041 CCTCTTCCGCTGATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA 2100  
Qy 2101 TTGTATGGGATGCTCTAAGGCTCATGATCTAAGCCTGGAAAGAAATATCTTACTGTG 2160  
Db 2101 TTGTATGGGATGCTCTAAGGCTCATGATCTAAGCCTGGAAAGAAATATCTTACTGTG 2160  
Qy 2161 CTGTTCCATATATGTTGG--TCTCAGGTGCAAGTTGGTGAATTAATCGGTTTAAAGGAGTCA 2218  
Db 2161 CTGTTCCATATATGTTGGTCTCTCTCAGGTGCAAGTTGGTGAATTAATCGGTTTAAAGGAGTCA 2220  
Qy 2219 AGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAACA 2278  
Db 2221 AGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAACA 2280  
Qy 2279 GGGGATCCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAA 2338  
Db 2281 GGGGATCCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAA 2340  
Qy 2339 TAGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTTCTCGATATGATTTCATG 2398  
Db 2341 TAGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTTCTCGATATGATTTCATG 2400  
Qy 2399 ACTTAGATCGTGGGCAATCCACGCTGCTCTATGAGGATACCTCTCCCTCATGGCAT 2458  
Db 2401 ACTTAGATCGTGGGCAATCCACGCTGCTCTATGAGGATACCTCTCCCTCATGGCAT 2460  
Qy 2459 TAATGAGAGGTGAGATATCTTCAAGGTTGCTATTTGCTGGGCGCCAGTCCTCTGGA 2518  
Db 2461 TAATGAGAGGTGAGATATCTTCAAGGTTGCTATTTGCTGGGCGCCAGTCCTCTGGA 2520  
Qy 2519 TCTTCTATGATACAGGATACAGGAACTTATATGGGTGACCTTGACCCAGCAAGTGAACAGG 2578  
Db 2521 TCTTCTATGATACAGGATACAGGAACTTATATGGGTGACCTTGACCCAGCAAGTGAACAGG 2580  
Qy 2579 GCTATTACTTAGGATCTGTGGCCATGCAAGCAAAAAGTTCCCTCTGAAACCAAAATCGTT 2638  
Db 2581 GCTATTACTTAGGATCTGTGGCCATGCAAGCAAAAAGTTCCCTCTGAAACCAAAATCGTT 2640  
Qy 2639 TACTGCTCTTACATGTTTTCTCGATGAGATGTCATTGTCATACACAGTATATTAC 2698  
Db 2641 TACTGCTCTTACATGTTTTCTCGATGAGATGTCATTGTCATACACAGTATATTAC 2700  
Qy 2699 TGAGTTTTTATGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGAGAC 2758  
Db 2701 TGAGTTTTTATGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGAGAC 2760  
Qy 2759 ACAGATAAGAGTTCTTGAATCGGAGAACTTATGAATGCAATCTTTTGGCATCTCTC 2818  
Db 2761 ACAGATAAGAGTTCTTGAATCGGAGAACTTATGAATGCAATCTTTTGGCATCTCTC 2820

QY 2819 AAGAAACCTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGAGA 2878  
DB 2821 AAGAAAAACCTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGAGA 2880  
QY 2879 ACTCTCTGTATACACCTGGCTATTAAACCAATGAGGAGGTTTAATCAACAGAAAAACACA 2938  
DB 2881 ACTCTCTGTATACACCTGGCTATTAAACCAATGAGGAGGTTTAATCAACAGAAAAACACA 2940  
QY 2939 GAATTGATCATCATATTTTGATACCTGCGCATGTAACATCTACTCTCGTGAATAAATGTGG 2998  
DB 2941 GAATTGATCATCATATTTTGATACCTGCGCATGTAACATCTACTCTCGTGAATAAATGTGG 3000  
QY 2999 TGCCATGACGGGCTACAGGTTTGTGTTAGTAACTTAATACCTTAACCCCATGCTCAA 3058  
DB 3001 TGCCATGACGGGCTACAGGTTTGTGTTAGTAACTTAATACCTTAACCCCATGCTCAA 3060  
QY 3059 AATCAATGATACATATTTCTGAGAGACCCAGCAATACCATGAAGTAATTAACAAAAAAA 3118  
DB 3061 AATCAATGATACATATTTCTGAGAGACCCAGCAATACCATGAAGTAATTAACAAAAAAA 3120  
QY 3119 AA 3120  
DB 3121 AA 3122  
  
RESULT 6  
ADV43981  
ID ADV43981 standard; cDNA; 2649 BP.  
AC ADV43981;  
XX  
DT 10-MAR-2005 (first entry)  
XX  
DE Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 1609.  
XX  
KW microarray; psychoneuroendocrinimmune; chronic fatigue;  
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;  
KW cancer; neoplasm; infection; expressed sequence tag; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO2004108899-A2.  
XX  
PD 16-DEC-2004.  
XX  
XX 04-JUN-2004; 2004WO-US017686.  
XX  
XX 04-JUN-2003; 2003US-0475915P.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Nicholson A, Vernon SD;  
XX  
XX WPI; 2005-031682/03.  
XX  
XX New microarray comprising probes for genes involved in  
PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a  
PT condition associated with PNI activity, e.g., inflammatory or infectious  
PT diseases.  
XX  
XX Claim 1; SEQ ID NO 1609; 254pp; English.  
XX  
XX The invention relates to a new microarray which comprises probes for  
CC genes involved in psychoneuroendocrinimmune (PNI) activity. The  
CC microarray is useful in diagnosing a condition associated with PNI  
CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,  
CC cancer and infection. The present sequence represents a  
CC psychoneuroendocrinimmune gene expressed sequence tag. Note the  
CC specification mentions SEQ ID NO of up to 3314 but only sequences up to  
CC SEQ ID NO 1829 are provided.  
XX  
XX Sequence 2649 BP; 804 A; 514 C; 585 G; 746 T; 0 U; 0 Other;

Query Match 83.3%; Score 2598; DB 14; Length 2649;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2648; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 214 ATGCGACGACCAATGGAAACAGAAACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 273  
DB 1 ATGCGACGACCAATGGAAACAGAAACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 60  
QY 274 GAGGAGAAATTTGAATCAGAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGGGGTAT 333  
DB 61 GAGGAGAAATTTGAATCAGAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGGGGTAT 120  
QY 334 TCCTGGATCAGCTTAAAGCTGCTTCGCGATACAGAAAAATATCATGGCTACATGATG 393  
DB 121 TCCTGGATCAGCTTAAAGCTGCTTCGCGATACAGAAAAATATCATGGCTACATGATG 180  
QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453  
DB 181 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 240  
QY 454 GACAGAAATCTATTACCTTGCCTGCTGCTGAGAAACAGAAAAATACACTGTTTATTTCT 513  
DB 241 GACAGAAATCTATTACCTTGCCTGCTGCTGAGAAACAGAAAAATACACTGTTTATTTCT 300  
QY 514 GAAATTCCTCAAACTATCAATAGAGCAGGATCTTAATGCTCTCTTGAAGCCCTTTTGTG 573  
DB 301 GAAATTCCTCAAACTATCAATAGAGCAGGATCTTAATGCTCTCTTGAAGCCCTTTTGTG 360  
QY 574 GATCTTTTTCAGGCACTGGAATGATTTCTCGAGAGAGAGAACTATTAGA 633  
DB 361 GATCTTTTTCAGGCACTGGAATGATTTCTCGAGAGAGAGAACTATTAGA 420  
QY 634 GAAAGAAACGCTTGAACAGCTGCGAATTTGCTTACGATTTATCACCAGGAAGTGA 693  
DB 421 GAAAGAAACGCTTGAACAGCTGCGAATTTGCTTACGATTTATCACCAGGAAGTGA 480  
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGAGTGGAGGGCCACAAGGA 753  
DB 481 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGAGTGGAGGGCCACAAGGA 540  
QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACCTAGTGTGCCCAATACGAGATG 813  
DB 541 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACCTAGTGTGCCCAATACGAGATG 600  
QY 814 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTATACATAGCAACGATAT 873  
DB 601 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTATACATAGCAACGATAT 660  
QY 874 TGGATATCTACATCGTAACACAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTA 933  
DB 661 TGGATATCTAACATCGTAACACAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTA 720  
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCCGTACCTTTGTTCTCCAAGAAGAA 993  
DB 721 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCCGTACCTTTGTTCTCCAAGAAGAA 780  
QY 994 TTTGATAGATATTTGCTATTGCTGTTTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 1053  
DB 781 TTTGATAGATATTTGCTATTGCTGTTTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 840  
QY 1054 AAAATTTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTT 1113  
DB 841 AAAATTTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTT 900  
QY 1114 ACATCCCTATGTTGGAACAGAGGGGCTAGATTTCCGTTATCCGTTATCCTAAAACAGGTACA 1173  
DB 901 ACATCCCTATGTTGGAACAGAGGGGCTAGATTTCCGTTATCCTAAAACAGGTACA 960  
QY 1174 GCATAATCTAAAGTCACCTTTTAAAGATCTCAGAAATTAATGATGATCTCAGAGGAAGATC 1233  
DB 961 GCATAATCTAAAGTCACCTTTTAAAGATCTCAGAAATTAATGATGATCTCAGAGGAAGATC 1020



QY 1234 ATAGATGTCATAGATAAGGAACAAATTCACCTTTTGGAGATTCCTATTTGAGGAGTTGAA 1293  
Db 1021 ATAGATGTCATAGATAAGGAACAAATTCACCTTTTGGAGATTCCTATTTGAGGAGTTGAA 1080  
QY 1294 TATATTGCGCAGAGTCGATGGACCTCTGAGGGAATATATGCTTGGTCCATCCTACTAGAT 1353  
Db 1081 TATATTGCGCAGAGTCGATGGACCTCTGAGGGAATATATGCTTGGTCCATCCTACTAGAT 1140  
QY 1354 CGCTCCAGAGTCGCTTACAGATAGTGTGATCTCACTGGAATTTATATCCAGTAGAA 1413  
Db 1141 CGCTCCAGAGTCGCTTACAGATAGTGTGATCTCACTGGAATTTATATCCAGTAGAA 1200  
QY 1414 GATGATGTTATGGAAGGCGAGAGCTCATTTGAGTCAGTGCCTGATCTCTGTGACGCCACTA 1473  
Db 1201 GATGATGTTATGGAAGGCGAGAGCTCATTTGAGTCAGTGCCTGATCTCTGTGACGCCACTA 1260  
QY 1474 ATTATCTATGAAGAAACAAACAGACATCTGTGATAAATATCCATGACATCTTTCATGTTTTT 1533  
Db 1261 ATTATCTATGAAGAAACAAACAGACATCTGTGATAAATATCCATGACATCTTTCATGTTTTT 1320  
QY 1534 CCCCAGAGTCACGAAGAGGAAATTTGAGTTTATTTTGGCTCTGGAATGCAAAACAGGTTTC 1593  
Db 1321 CCCCAGAGTCACGAAGAGGAAATTTGAGTTTATTTTGGCTCTGGAATGCAAAACAGGTTTC 1380  
QY 1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT 1653  
Db 1381 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT 1440  
QY 1654 GGGCTGCTCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATGCAATTTACCACT 1713  
Db 1441 GGGCTGCTCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATGCAATTTACCACT 1500  
QY 1714 GGTGAATGGGAAGTCTTGGCGGCGATGATCTATATATCAAGTTGATGAAGTCAGAAG 1773  
Db 1501 GGTGAATGGGAAGTCTTGGCGGCGATGATCTATATCAAGTTGATGAAGTCAGAAG 1560  
QY 1774 CTGCTATATTTGAGAGCCAAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT 1833  
Db 1561 CTGCTATATTTGAGAGCCAAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT 1620  
QY 1834 TACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893  
Db 1621 TACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1680  
QY 1894 ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAAACAGAGAAATCCACACTGT 1953  
Db 1681 ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAAACAGAGAAATCCACACTGT 1740  
QY 1954 GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAACAAAGGAATTT 2013  
Db 1741 GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAACAAAGGAATTT 1800  
QY 2014 TGGGCCACCAATTTTGGATTCAGCAGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073  
Db 1801 TGGGCCACCAATTTTGGATTCAGCAGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 1860  
QY 2074 TCTTTTGAAGTACTACTCGATTTACATTTGATGATGCTCTCAAGCTCTCATGATCTA 2133  
Db 1861 TCTTTTGAAGTACTACTCGATTTACATTTGATGATGCTCTCAAGCTCTCATGATCTA 1920  
QY 2134 CAGCTGGAAGAAATATCTCTCTGCTGTTCTATATATGTTGCTCTCAGGTGCAAGTTG 2193  
Db 1921 CAGCTGGAAGAAATATCTCTCTGCTGTTCTATATATGTTGCTCTCAGGTGCAAGTTG 1980  
QY 2194 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGT 2253  
Db 1981 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGT 2040  
QY 2254 TATGTGGTTGTAGTAGACAAACAGGGGATCCCTGTCAACGAGGGCTTAAATTTGAAGGC 2313  
Db 2041 TATGTGGTTGTAGTAGACAAACAGGGGATCCCTGTCAACGAGGGCTTAAATTTGAAGGC 2100  
QY 2314 GCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT 2373

Db 2101 GCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT 2160  
QY 2374 CTAGCTTCTCGATATGATTTTCATTTAGATCGTGTGGGCATCCACGGCTGTCCTAT 2433  
Db 2161 CTAGCTTCTCGATATGATTTTCATTTAGATCGTGTGGGCATCCACGGCTGTCCTAT 2220  
QY 2434 GGAGGATACCTCTCCCTCGATGGCATTAATGCGAGAGTCCAGATATCTTTCAGGGTTCTCTATT 2493  
Db 2221 GGAGGATACCTCTCCCTCGATGGCATTAATGCGAGAGTCCAGATATCTTTCAGGGTTCTCTATT 2280  
QY 2494 GCTGGGGCCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553  
Db 2281 GCTGGGGCCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2340  
QY 2554 GGTCAACCTCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613  
Db 2341 GGTCAACCTCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2400  
QY 2614 AAGTTTCCCTCTGAAACCAATCGTTTACTGCTTACATGCTTACATGCTTCTCGATGAGNATGTC 2673  
Db 2401 AAGTTTCCCTCTGAAACCAATCGTTTACTGCTTACATGCTTCTCGATGAGNATGTC 2460  
QY 2674 CATTTTGCACATACCACTATATTAAGTCTTACTGAGTCTTACTGAGGCTGGAAGGCCATATGAT 2733  
Db 2461 CATTTTGCACATACCACTATATTAAGTCTTACTGAGTCTTACTGAGGCTGGAAGGCCATATGAT 2520  
QY 2734 TTACAGATCTATCTCTCAGAGAGACACAGCATAAAGTTCCTGAAATCGGAGAAACATATAT 2793  
Db 2521 TTACAGATCTATCTCTCAGAGAGACACAGCATAAAGTTCCTGAAATCGGAGAAACATATAT 2580  
QY 2794 GAACCTGATCTTTTGGTCACTACCTTCAAGAAAAACCTTGGATCACGTTATGCTGCTCTAAAA 2853  
Db 2581 GAACCTGATCTTTTGGTCACTACCTTCAAGAAAAACCTTGGATCACGTTATGCTGCTCTAAAA 2640  
QY 2854 GTGATATAA 2862  
Db 2641 GTGATATAA 2649  
RESULT 7  
AAH99935  
ID AAH99935 standard; cDNA; 2643 BP.  
XX AAH99935;  
XX  
XX  
DT 12-APR-2002 (first entry)  
XX  
DE Coding sequence of 21953 human prollyl oligopeptidase.  
XX  
KW 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;  
KW cardiovascular disease; autoimmune disease; atopic allergy;  
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;  
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;  
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;  
KW Grave's disease; neuronal disorder; demyelinating disease; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200179473-A2.  
XX  
XX  
PD 25-OCT-2001.  
XX  
PF 11-APR-2001; 2001WO-US040483.  
XX  
PR 18-APR-2000; 2000US-0197508P.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Meyers RA, Williamson M;  
XX  
DR WPI; 2002-034353/04.  
DR P-PSDB; AAG78415.

XX New polypeptides 21953, member of human prollyl oligopeptidase family,  
PT useful as diagnostic targets and therapeutic agents for controlling  
PT cancer, lymphoma and leukemia.  
XX  
XX  
PS Claim 2; Page; 121pp; English.  
XX  
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.  
CC Which is cytostatic, antidiabetic, antithrombotic, neuroprotective,  
CC antihypertensive, antiproliferative, antiparasitic, antitubercular,  
CC ophthalmological, antiinflammatory, nootropic, antiparkinsonian,  
CC anticonvulsant, gynaecological, vasotropic, antianginal, cardiant,  
CC antithrombotic, anorectic and metabolic in its action. Uses include  
CC gene therapy, expression or activity of 21953 protein modulator, it is  
CC useful for identifying a compound which binds to it and can be used in  
CC preventing, treating or detecting a cellular proliferative or  
CC differentiative disorder. The 21953 molecules can act as novel diagnostic  
CC targets and therapeutic agents for controlling disorders associated with  
CC the aberrant activity or degradation of peptide hormones such as cancer,  
CC associated with cell differentiation and proliferation such as cancer,  
CC immune function, reproductive, neurological and cardiovascular function.  
CC The 21953 molecules are thus useful for treating and preventing cellular  
CC proliferative and differentiative disorders, haematopoietic neoplastic  
CC disorders, immune disorders such as autoimmune diseases, diabetes  
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,  
CC neuronal disorders, demyelinating diseases, vascular disorders and  
CC metabolism or pain disorders. This sequence represents the cDNA encoding  
CC sequence of 21953 human prollyl oligopeptidase. This sequence represents  
CC the coding sequence of 21953, being the sequence in between the start and  
CC the stop codon of the sequence represented in AH99934  
XX  
SQ Sequence 2643 BP; 800 A; 514 C; 585 G; 744 T; 0 U; 0 Other;

Query Match 81.4%; Score 2541; DB 6; Length 2643;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2641; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

217 GCAGCAGCAATGAAACAGAACAGCTGGGTGTGAGATATTGAACTGGGCTGTGAG 276  
1 GCAGCAGCAATGAAACAGAACAGCTGGGTGTGAGATATTGAACTGGGCTGTGAG 60  
277 GAGAAATTTGAATCAGAGTGGCTTAAATTTGAGCGCTTTTATTTGAGCGGTATTC 336  
61 GAGAAATTTGAATCAGAGTGGCTTAAATTTGAGCGCTTTTATTTGAGCGGTATTC 120  
337 TGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGGCTACATGGCT 396  
121 TGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGGCTACATGGCT 180  
397 AAGCACCACATGATTTCAATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAG 456  
181 AAGCACCACATGATTTCAATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAG 240  
457 AGAATCTATTACCTTGGCCATGCTGGTGAGAACAGAAATACACTGTTTATTCGAA 516  
241 AGAATCTATTACCTTGGCCATGCTGGTGAGAACAGAAATACACTGTTTATTCGAA 300  
517 ATTCCTCAAAATCATCAATAGCAGCAGCTTAAATGCTCTCTTGAAGCCTCTTTGGAT 576  
301 ATTCCTCAAAATCATCAATAGCAGCAGCTTAAATGCTCTCTTGAAGCCTCTTTGGAT 360  
577 CTTTTTCAGGCAACCTGGATATGGAATGATTTCTCGAGAAAGAACTATTAAAGAGAA 636  
361 CTTTTTCAGGCAACCTGGATATGGAATGATTTCTCGAGAAAGAACTATTAAAGAGAA 420  
637 AGAATCCCATTTGAAACAGTCCGAATTTCTTACGATTATCAACAAGGAAGTGAACA 696  
421 AGAATCCCATTTGAAACAGTCCGAATTTCTTACGATTATCAACAAGGAAGTGAACA 480  
697 TTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAAAGATCGAGGGCCACAAGGATTT 756  
481 TTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAAAGATCGAGGGCCACAAGGATTT 540

QY 757 ACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATGGAT 816  
DB 541 ACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATGGAT 600  
QY 817 CCAAAATTTATGCCCGCTGATCCAGACTGCGATTTTATACATAGCAACATATTTGG 876  
DB 601 CAAAATTTATGCCCGCTGATCCAGACTGCGATTTTATACATAGCAACATATTTGG 660  
QY 877 ATATCTAACTGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACATAGCTAGCC 936  
DB 661 ATATCTAACTGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACATAGCTAGCC 720  
QY 937 AACATGGAAGAGATGCCAGATCAGCTGGAGTGGCTTCTTCTCAAGAGAAATTT 996  
DB 721 AACATGGAAGAGATGCCAGATCAGCTGGAGTGGCTTCTTCTCAAGAGAAATTT 780  
QY 997 GATAGATTTCTGGCTATTTGGTGTCCAAAAGCTGAAAACAACTCCCAAGTGGTGTAAA 1056  
DB 781 GATAGATTTCTGGCTATTTGGTGTCCAAAAGCTGAAAACAACTCCCAAGTGGTGTAAA 840  
QY 1057 ATTCTTATGAAATTTCTATATGAAAGAAATGATGAATCTGAGGTGAAATTTATTCATGTTACA 1116  
DB 841 ATTCTTATGAAATTTCTATATGAAAGAAATGATGAATCTGAGGTGAAATTTATTCATGTTACA 900  
QY 1117 TCCCTTATCTGGAACCAAGGAGGAGGAGATTCATCTCCGTTATCTCTAAACAGGTACAGCA 1176  
DB 901 TCCCTTATCTGGAACCAAGGAGGAGGAGATTCATCTCCGTTATCTCTAAACAGGTACAGCA 960  
QY 1177 AATCTTAAAGTCACTTTTAAAGTGTGCAAAATATGATGCTGAAAGAGAGATCATATA 1236  
DB 961 AATCTTAAAGTCACTTTTAAAGTGTGCAAAATATGATGCTGAAAGAGAGATCATATA 1020  
QY 1237 GATGTCATAGATAAGGAACCTAATTCACCTTTTGGAGATTTCTATTTGAAGAGTGAATAT 1296  
DB 1021 GATGTCATAGATAAGGAACCTAATTCACCTTTTGGAGATTTCTATTTGAAGAGTGAATAT 1080  
QY 1297 ATTGCCAGAGCTGGATGGAATCTCTGAGGAAATATGCTGGTCCATCTCTACTAGATCGC 1356  
DB 1081 ATTGCCAGAGCTGGATGGAATCTCTGAGGAAATATGCTGGTCCATCTCTACTAGATCGC 1140  
QY 1357 TCCAGACTGCTCAGATAGTGTGATCTCAGCTGAATTTATTTATCCAGATGAGAAGAT 1416  
DB 1141 TCCAGACTGCTCAGATAGTGTGATCTCAGCTGAATTTATTTATCCAGTGAAGAT 1200  
QY 1417 GATGTTATGAAAGGAGAGAGACTCATTTGAGTCACTGCTGATTTCTGTGAGCCCACTAAT 1476  
DB 1201 GATGTTATGAAAGGAGAGAGACTCATTTGAGTCACTGCTGATTTCTGTGAGCCCACTAAT 1260  
QY 1477 ATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTATGTTTTCCC 1536  
DB 1261 ATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTATGTTTTCCC 1320  
QY 1537 CAAAGTCAAGAGAGAAATTTGAGTTATTTTGGCTCTGATGCAATCAAAACAGGTTCCTG 1596  
DB 1321 CAAAGTCAAGAGAGAAATTTGAGTTATTTTGGCTCTGATGCAATCAAAACAGGTTCCTG 1380  
QY 1597 CATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGTGG 1656  
DB 1381 CATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGTGG 1440  
QY 1657 CTGCTCTCTCAAGTGTCTATCAAGAGAGAGATAGCAATTTACCAAGTGGT 1716  
DB 1441 CTGCTCTCTCAAGTGTCTATCAAGAGAGAGATAGCAATTTACCAAGTGGT 1500  
QY 1717 GAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAGTGTGATGAAAGTGAAGGCTG 1776  
DB 1501 GAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAGTGTGATGAAAGTGAAGGCTG 1560  
QY 1777 GTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACTGTAGCTAGTTCAGTTTAC 1836  
DB 1561 GTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACTGTAGCTAGTTCAGTTTAC 1620  
QY 1837 GTAAATCTCGAGAGAGTGAACAGGCTGACTGACCCGGCTACTACATTTCTTGTGCTGATC 1896



Matches 2646; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	214	ATGGCAGCAGCAATCGAAACAGAACACAGCTGGGTGTTGAGATATTTGAAACTCGGACTGT	273
Db	1	ATGGCAGCAGCAATCGAAACAGAACACAGCTGGGTGTTGAGATATTTGAAACTCGGACTGT	60
QY	274	GAGGAGATATTTGAATCACAGGATCGGCCTAAATTTGGAGCTTTTATTTGTTGAGCGGTAT	333
Db	61	GAGGAGATATTTGAATCACAGGATCGGCCTAAATTTGGAGCTTTTATTTGTTGAGCGGTAT	120
QY	334	TCCTGGAGTCAGCTTAAAGCTTGGCCGATACCGATACAGAAATATCATGGCTACATGATG	393
Db	121	TCCTGGAGTCAGCTTAAAGCTTGGCCGATACCGATACAGAAATATCATGGCTACATGATG	180
QY	394	GCTAAGGCACACATGATTTCACTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA	453
Db	181	GCTAAGGCACACATGATTTCACTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA	240
QY	454	GACAGAACTATTAACCTTGGCATGCTCTGGTGAGAACAGAGAAATACACTGTTTATTTCT	513
Db	241	GACAGAACTATTAACCTTGGCATGCTCTGGTGAGAACAGAGAAATACACTGTTTATTTCT	300
QY	514	GAATTTCCCAAAATATCAATAGAGCAGCAGTCTTAATGCTCTCTTGAAGCCTCTTTTG	573
Db	301	GAATTTCCCAAAATATCAATAGAGCAGCAGTCTTAATGCTCTCTTGAAGCCTCTTTTG	360
QY	574	GATCTTTTTCAGGCAACACATGGAATATGTAATCTCGAGAGAAAGAACTATTAAAGA	633
Db	361	GATCTTTTTCAGGCAACACATGGAATATGTAATCTCGAGAGAAAGAACTATTAAAGA	420
QY	634	GAAGAAACGATTTGGAACACATCGGAATTTGCTTTACGATTTATCAACCAAGGAAGTGA	693
Db	421	GAAGAAACGATTTGGAACACATCGGAATTTGCTTTACGATTTATCAACCAAGGAAGTGA	480
QY	694	ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCGTAAAGATGGAGGGCCACAAGGA	753
Db	481	ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCGTAAAGATGGAGGGCCACAAGGA	540
QY	754	TTTAGGCAACACCTTTAAAGCCCATCTAGTGGAACTAGTGTGTCCTCAACATACGGATG	813
Db	541	TTTAGGCAACACCTTTAAAGCCCATCTAGTGGAACTAGTGTGTCCTCAACATACGGATG	600
QY	814	GATCCAAATATATGCCCGCTGATCCAGACTGATTTGCTTTATACATAGCAACGATATT	873
Db	601	GATCCAAATATATGCCCGCTGATCCAGACTGATTTGCTTTATACATAGCAACGATATT	660
QY	874	TGGATATCTAATCATCGTAAACAGAGAAAGAGAGACTCACTTATGTGCAAAATGAGCTA	933
Db	661	TGGATATCTAATCATCGTAAACAGAGAAAGAGAGACTCACTTATGTGCAAAATGAGCTA	720
QY	934	GCACATCGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGAA	993
Db	721	GCACATCGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGAA	780
QY	994	TTTGATAGATATCTGGCTATTTGGTGTGTCCTGAGCTGAGTAAACCTCCAGAGTGGT	1053
Db	781	TTTGATAGATATCTGGCTATTTGGTGTGTCCTGAGCTGAGTAAACCTCCAGAGTGGT	840
QY	1054	AAATTTCTTGAATTTATATGAAAGAAATGATGATCTGAGGTGGAAATTTATTCATGTT	1113
Db	841	AAATTTCTTGAATTTATATGAAAGAAATGATGATCTGAGGTGGAAATTTATTCATGTT	900
QY	1114	AdATCCCTATGTTGGAACAAAGAGGGCAGATTTCTTCGGTTATCTTAAACAGGTACA	1173
Db	901	AdATCCCTATGTTGGAACAAAGAGGGCAGATTTCTTCGGTTATCTTAAACAGGTACA	960
QY	1174	GCAAATCTTAAAGTCATTTTAAGATGTCAGAAATATGATTTGATGCTGGAAGAGGATC	1233
Db	961	GCAAATCTTAAAGTCATTTTAAGATGTCAGAAATATGATTTGATGCTGGAAGAGGATC	1020
QY	1234	ATAGATGTCATAGATAAGGAACATTAATCAACCTTTTGAAGTCTTATTTGAAGAGTTGAA	1293
Db	1021	ATAGATGTCATAGATAAGGAACATTAATCAACCTTTTGAAGTCTTATTTGAAGAGTTGAA	1080

QY	1294	TATATTTCCAGAGCTGGATGGAATCTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1353
Db	1081	TATATTTCCAGAGCTGGATGGAATCTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1140
QY	1354	CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACTGGAATTTATTTCCAGTAGAA	1413
Db	1141	CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACTGGAATTTATTTCCAGTAGAA	1200
QY	1414	GATGATGTTATGGAAAGGACAGACTCAATGAGTCACTGCTGATCTCTGAGCGCACTA	1473
Db	1201	GATGATGTTATGGAAAGGACAGACTCAATGAGTCACTGCTGATCTCTGAGCGCACTA	1260
QY	1474	ATTATCTATGAAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1533
Db	1261	ATTATCTATGAAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1320
QY	1534	CCCCAAGTCAAGAAAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAAACAGTTTC	1593
Db	1321	CCCCAAGTCAAGAAAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAAACAGTTTC	1380
QY	1594	CGTCATTTTATCAAAATTTACATCTATTTTAAAGAAAGCAAAATATAACGATCCAGTGT	1653
Db	1381	CGTCATTTTATCAAAATTTACATCTATTTTAAAGAAAGCAAAATATAACGATCCAGTGT	1440
QY	1654	GGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAAAGAGGAGATAGCAATACCACT	1713
Db	1441	GGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAAAGAGGAGATAGCAATACCACT	1500
QY	1714	GGTGATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAGTCAAGAG	1773
Db	1501	GGTGATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAGTCAAGAG	1560
QY	1774	CTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACTGTGACGTAGTCACT	1833
Db	1561	CTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACTGTGACGTAGTCACT	1620
QY	1834	TACGTAAATCTCGAGAGGTGACAAGCTGACGCGTGGCTACTCACATTTCTTGCTGC	1893
Db	1621	TACGTAAATCTCGAGAGGTGACAAGCTGACGCGTGGCTACTCACATTTCTTGCTGC	1680
QY	1894	ATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGT	1953
Db	1681	ATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGT	1740
QY	1954	GTGTCCCTTTAAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAACAAAGGAAAT	2013
Db	1741	GTGTCCCTTTAAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAACAAAGGAAAT	1800
QY	2014	TGGGCGCACATTTTGGATTCAGCAGGTCTCTTCTCTGACTATATCTCTCAGAAAATTTTC	2073
Db	1801	TGGGCGCACATTTTGGATTCAGCAGGTCTCTTCTCTGACTATATCTCTCAGAAAATTTTC	1860
QY	2074	TCCTTTGAAAGTACTACTGGATTTTACATGTATGGGATGCTCTCAAGCCTCATGATCTA	2133
Db	1861	TCCTTTGAAAGTACTACTGGATTTTACATGTATGGGATGCTCTCAAGCCTCATGATCTA	1920
QY	2134	CAGCTGGAAGAAATATCTTACTGTCTTATATATGTTGGTCTCTCAGGTGAGTTG	2193
Db	1921	CAGCTGGAAGAAATATCTTACTGTCTTATATATGTTGGTCTCTCAGGTGAGTTG	1980
QY	2194	GTGAATATCTGTTTAAAGGAGTCAAGTATTTTCCGCTTGAATACCTAGCCTCTCTAGGT	2253
Db	1981	GTGAATATCTGTTTAAAGGAGTCAAGTATTTTCCGCTTGAATACCTAGCCTCTCTAGGT	2040
QY	2254	TATGTGTTGTAGTAGAGCAACAGGGATCTGTCAACCGAGGCTTAAATTTGAAGGC	2313
Db	2041	TATGTGTTGTAGTAGAGCAACAGGGATCTGTCAACCGAGGCTTAAATTTGAAGGC	2100
QY	2314	GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACATCAGGTGGAAGGATCCAAAT	2373
Db	2101	GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACATCAGGTGGAAGGATCCAAAT	2160

QY 2374 CTAGCTTCTCGATATGATTTTCATTCAGTCTAGATCGCTGGGATCCAGCGGTGCTCTAT 2433  
Db |||||  
QY 2434 GGAGATACCTTCCTGTATGGCATTATTCAGAGGTGAGATATCTTCAGGGTGCCTATT 2493  
Db |||||  
QY 2493 GGAGATACCTTCCTGTATGGCATTATTCAGAGGTGAGATATCTTCAGGGTGCCTATT 2553  
Db |||||  
QY 2553 GCTGGGGCCCGCAGTCTCTGTGTGATCTTCATGATACAGGATACACGGAACGTTATATG 2613  
Db |||||  
QY 2613 GCTGGGGCCCGCAGTCTCTGTGTGATCTTCATGATACAGGATACACGGAACGTTATATG 2673  
Db |||||  
QY 2673 GCTGGGGCCCGCAGTCTCTGTGTGATCTTCATGATACAGGATACACGGAACGTTATATG 2733  
Db |||||  
QY 2733 GCTGGGGCCCGCAGTCTCTGTGTGATCTTCATGATACAGGATACACGGAACGTTATATG 2793  
Db |||||  
QY 2793 GCTGGGGCCCGCAGTCTCTGTGTGATCTTCATGATACAGGATACACGGAACGTTATATG 2853  
Db |||||  
QY 2853 GCTGGGGCCCGCAGTCTCTGTGTGATCTTCATGATACAGGATACACGGAACGTTATATG 2913  
Db |||||  
QY 2913 GCTGGGGCCCGCAGTCTCTGTGTGATCTTCATGATACAGGATACACGGAACGTTATATG 2973  
Db |||||  
QY 2973 GCTGGGGCCCGCAGTCTCTGTGTGATCTTCATGATACAGGATACACGGAACGTTATATG 3033  
Db |||||

## RESULT 9

ID ABK83322 standard; cDNA; 2671 BP.  
XX  
AC ABK83322;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE cDNA encoding human DPP4V related serine protease DPP-1.  
XX  
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPP; DPP4V; infection; human immunodeficiency virus; HIV-1; HIV-2; pain; diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke; heart failure; hypertension; urinary retention; osteoporosis; cancer; ulcer; allergy; cancer; psychotic disorder; neurological disorder; dyskinesia; reproductive disorder; inflammatory disorder; metabolic disorder; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200231134-A2.  
XX  
PD 18-APR-2002.  
XX  
PF 12-OCT-2001; 2001WO-US031874.  
XX  
PR 12-OCT-2000; 2000US-0240117P.  
XX  
PA (FERR ) FERRING BV.  
XX  
PI Qi S, Akinsanya KO, Riviere PJ, Junien J;  
XX  
DR WPI; 2002-444178/P.  
XX  
DR P-PSDB; ABG61591.  
XX  
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding

PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and viral infections, cancers, allergies, neurological disorders, or pain.  
PS Claim 1; Page 53-54; 113pp; English.  
XX The present invention relates to the isolation of novel human serine proteases referred to as dipeptidyl peptidase IV (DPP4V)-related proteins (DPP4V). The dipeptidyl peptidase IV-related proteins (DPP4V) and nucleic acids encoding them are useful for treating infections such as fungal, bacterial, protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes, precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, psychotic and neurodegenerative disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias. These may also be used in discovering therapeutic agents for the treatment of reproductive, inflammatory and metabolic disorders. ABK83322 -ABK83343 encode human DPP4V proteins  
XX  
SQ Sequence 2671 BP; 805 A; 524 C; 594 G; 748 T; 0 U; 0 Other;  
Query Match 80.0%; Score 2495; DB 6; Length 2671;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 213 CATGGCGAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAACTGGGACTG 272  
Db |||||  
QY 7 CATGGCGAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAACTGGGACTG 66  
Db |||||  
QY 273 TGAGGAAATATTGAATCACAGGATCGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTA 332  
Db |||||  
QY 67 TGAGGAAATATTGAATCACAGGATCGGCTAAATTTGGAGCGCTTTTATGTTGAGCGGTA 126  
Db |||||  
QY 333 TTCCTGGAGTCAGCTTAAAGAGCTGCTCCGATACCGAATAATATCATGCTACATGAT 392  
Db |||||  
QY 127 TTCCTGGAGTCAGCTTAAAGAGCTGCTCCGATACCGAATAATATCATGCTACATGAT 186  
Db |||||  
QY 393 GGCTAAGGCGACCATGATTTTCAATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 452  
Db |||||  
QY 187 GGCTAAGGCGACCATGATTTTCAATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 246  
Db |||||  
QY 453 AGACAGAAATCTATTACCTTGGCATGCTGTGTGAGAAACAGAGAAATAACATCTGTTTATTC 512  
Db |||||  
QY 247 AGACAGAAATCTATTACCTTGGCATGCTGTGTGAGAAACAGAGAAATAACATCTGTTTATTC 306  
Db |||||  
QY 513 TGAATTTCCCAAACTATCAATAGACAGCTCTTAATGCTCTCTTGGAGCGCTCTTTT 572  
Db |||||  
QY 307 TGAATTTCCCAAACTATCAATAGACAGCTCTTAATGCTCTCTTGGAGCGCTCTTTT 366  
QY 573 GGATCTTTTTCAGGCAACACTGGACTATGATGATTTCTCGAGAGAGAACTATTAAAG 632  
Db |||||  
QY 367 GGATCTTTTTCAGGCAACACTGGACTATGATGATTTCTCGAGAGAGAACTATTAAAG 426  
Db |||||  
QY 633 AGAAGAAACCGATTTGGAAACAGTCCGATTCCTTACGATTTATCACAGAGAGCTGG 692  
Db |||||  
QY 427 AGAAGAAACCGATTTGGAAACAGTCCGATTCCTTACGATTTATCACAGAGAGCTGG 486  
Db |||||  
QY 693 AACATTTCTGTTTCAAGCGGTAGTGAATTTATCACGTAAAGATGAGGGGCCACAGG 752  
Db |||||  
QY 487 AACATTTCTGTTTCAAGCGGTAGTGAATTTATCACGTAAAGATGAGGGGCCACAGG 546  
Db |||||  
QY 753 ATTTACCGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTGTGTCCTCAACATACCGAT 812  
Db |||||  
QY 547 ATTTACCGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTGTGTCCTCAACATACCGAT 606  
Db |||||  
QY 813 GGATCCAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 872  
Db |||||  
QY 607 GGATCCAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 666  
Db |||||  
QY 873 TTGATATCTAACATCGTAAACAGAGAGAGAGAGACTCACTTATGTCACAACTAGACT 932  
Db |||||  
QY 667 TTGATATCTAACATCGTAAACAGAGAGAGAGAGACTCACTTATGTCACAACTAGACT 726  
Db |||||

```
QY 933 AGCAACATGGGAAGAGATGCCAGATCAGCTGGAGTCGCTACTCTTGGTCTCCAGAAGA 992
Db 727 AGCAACATGGGAAGAGATGCCAGATCAGCTGGAGTCGCTACTCTTGGTCTCCAGAAGA 786
QY 993 ATTTGATAGATATTCCTGGCTATTTGGTGTGCTCCAAAGCTGAAACAACTCCCAAGTGGTG 1052
Db 787 ATTTGATAGATATTCCTGGCTATTTGGTGTGCTCCAAAGCTGAAACAACTCCCAAGTGGTG 846
QY 1053 TAAAAATCTTAGAATCTATATCAAGAAAAATGATGAATCTGAGTGGAAAAATTAATCATGT 1112
Db 847 TAAAAATCTTAGAATCTATATGAAGAAAAATGATGAATCTGAGTGGAAAAATTAATCATGT 906
QY 1113 TACATCCCTATGTTGGAAACAAGGAGGCGAGATTCATTCGGTTATCTTAAACAGGTAC 1172
Db 907 TACATCCCTATGTTGGAAACAAGGAGGCGAGATTCATTCGGTTATCTTAAACAGGTAC 966
QY 1173 AGCAATCTTAAAGTCACCTTTTAAAGATGCAGAAATATGATGATGCTGAAGGAAGAT 1232
Db 967 AGCAATCTTAAAGTCACCTTTTAAAGATGCAGAAATATGATGATGCTGAAGGAAGAT 1026
QY 1233 CATAGATGTCATAGATAAGGAACTAATTCACCTTTTGAAGGAGTTGA 1292
Db 1027 CATAGATGTCATAGATAAGGAACTAATTCACCTTTTGAAGGAGTTGA 1086
QY 1293 ATATATGTCAGAGCTGGATGGACTCCTGAGGAAAAATATGCTTGGTCCATCTTACTAGA 1352
Db 1087 ATATATGTCAGAGCTGGATGGACTCCTGAGGAAAAATATGCTTGGTCCATCTTACTAGA 1146
QY 1353 TCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTAATTTATCCAGPAGA 1412
Db 1147 TCGCTCCAGACTCGCTCGAGATAGTGTGATCTCACCTGAAATTAATTTATCCAGPAGA 1206
QY 1413 AGATAGTGTATGGAAGGCGAGAGACTCAATGAGTCAGTGCCTGATTCGTGAGCGCACT 1472
Db 1207 AGATAGTGTATGGAAGGCGAGAGACTCAATGAGTCAGTGCCTGATTCGTGAGCGCACT 1266
QY 1473 AATTATCTATGAGAAAAACAAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT 1532
Db 1267 AATTATCTATGAGAAAAACAAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT 1326
QY 1533 TCCCAAGTCACGAGAGAAATTCAGTTTATTTTGGCTCTGAATCCAAAAACAGTTTT 1592
Db 1327 TCCCAAGTCACGAGAGAAATTCAGTTTATTTTGGCTCTGAATCCAAAAACAGTTTT 1386
QY 1593 CCGTCATTTATACAAAATACATCTATTTTAAAGGAAAGCAATAATAAACGATCCAGTGG 1652
Db 1387 CCGTCATTTATACAAAATACATCTATTTTAAAGGAAAGCAATAATAAACGATCCAGTGG 1446
QY 1653 TGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTAACAG 1712
Db 1447 TGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTAACAG 1506
QY 1713 TGGTGAATPGGAAGTTCTTGGCGGAGATGATCTAATATCCAAAGTGTGATGAGTCAGAG 1772
Db 1507 TGGTGAATPGGAAGTTCTTGGCGGAGATGATCTAATATCCAAAGTGTGATGAGTCAGAG 1566
QY 1773 GCTGGTATATTTTGAAGGCCAAAGACTCCCTTTTGAAGCATCACCTGTACGTAGTCAG 1832
Db 1567 GCTGGTATATTTTGAAGGCCAAAGACTCCCTTTTGAAGCATCACCTGTACGTAGTCAG 1626
QY 1833 TTACGTAATCTCGAGAGGTGACAAGGCTGACCTGCTGCTACTCACATTTCTTGTCTG 1892
Db 1627 TTACGTAATCTCGAGAGGTGACAAGGCTGACCTGCTGCTACTCACATTTCTTGTCTG 1686
QY 1893 CATCAGTCAGACTGTGACTTTCTTTAATAAGTAAGTATAGTAACCAAGAAATCCACACTG 1952
Db 1687 CATCAGTCAGACTGTGACTTTCTTTAATAAGTATAGTATACCAAGAAATCCACACTG 1746
QY 1953 TGTGTCCCTTTTACAGCTATCAAGTCTGGAAGATGCCAACTTGGAAAAACAAGGAATTT 2012
Db 1747 TGTGTCCCTTTTACAGCTATCAAGTCTGGAAGATGCCAACTTGGAAAAACAAGGAATTT 1806
```

```
QY 2013 TTGGGCCACCAATTTTGGATTCCAGAGGTCTCTTTCTTGACTATATCTCTCCAGAAATTTT 2072
Db 1807 TTGGGCCACCAATTTTGGATTCCAGAGGTCTCTTTCTTGACTATATCTCTCCAGAAATTTT 1866
QY 2073 CTCTTTTGAAGTACTACTCGAATTTACATGTATGGGATGCTCTACAAGCCTCATGATCT 2132
Db 1867 CTCTTTTGAAGTACTACTCGAATTTACATGTATGGGATGCTCTACAAGCCTCATGATCT 1926
QY 2133 ACAGCCTGGAAGAATAATCTCTACTGCTCTTATATATGGTGGTCTCTCAGGTGCAAGTT 2192
Db 1927 ACAGCCTGGAAGAATAATCTCTACTGCTCTTATATATGGTGGTCTCTCAGGTGCAAGTT 1986
QY 2193 GGTGAATAATTCGGTTTAAAGAGTCAAGTATTTTCCGTTTGAATACCCCTAGCCTCTCTAGG 2252
Db 1987 GGTGAATAATTCGGTTTAAAGAGTCAAGTATTTTCCGTTTGAATACCCCTAGCCTCTCTAGG 2046
QY 2253 TTATGTGGTTGTAGTGATAGACAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTGAAGG 2312
Db 2047 TTATGTGGTTGTAGTGATAGACAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTGAAGG 2106
QY 2313 CGCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGACTCCAAATA 2372
Db 2107 CGCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGACTCCAAATA 2166
QY 2373 TCTAGCTTCTCGATATGATTTCAITTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTA 2432
Db 2167 TCTAGCTTCTCGATATGATTTCAITTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTA 2226
QY 2433 TGGAGGATACCTCTCCCTCATGCGCATTAATGCAGAGGTCCAGATATCTTCAGGGTTCCTAT 2492
Db 2227 TGGAGGATACCTCTCCCTCATGCGCATTAATGCAGAGGTCCAGATATCTTCAGGGTTCCTAT 2286
QY 2493 TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATAT 2552
Db 2287 TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATAT 2346
QY 2553 GGGTCACCCCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA 2612
Db 2347 GGGTCACCCCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGA 2406
QY 2613 AAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAATGT 2672
Db 2407 AAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAATGT 2466
QY 2673 CCATTTTGCACATACCAAGTATATTAATGATTTTTTGTAGGGCTGGAAGCCATATGA 2732
Db 2467 CCATTTTGCACATACCAAGTATATTAATGATTTTTTGTAGGGCTGGAAGCCATATGA 2526
QY 2733 TTTACAGATCTATCTCTCAGGAGACACAGCATAAAGAGTTCTGAATCGGAGAACATTA 2792
Db 2527 TTTACAGATCTATCTCTCAGGAGACACAGCATAAAGAGTTCTGAATCGGAGAACATTA 2586
QY 2793 TGAACCTGCATCTTTTGCACCTACCTTCAAGAAAAACCTTGGATCAAGTATGCTGCTTAAA 2852
Db 2587 TGAACCTGCATCTTTTGCACCTACCTTCAAGAAAAACCTTGGATCAAGTATGCTGCTTAAA 2646
QY 2853 AGTGATAT 2860
Db 2647 AGTGATAT 2654
```

RESULT 10  
ADL13374

ID ADL13374 standard; cDNA; 2797 BP.

XX ADL13374;

XX 06-MAY-2004 (first entry)

DE Human steroid-induced C3A liver cell cDNA #1103.

XX ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;  
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.



XX Homo sapiens.  
XX US6673549-B1.  
XX PD 06-JAN-2004.  
XX PF 12-OCT-2001; 2001US-00976594.  
XX PR 12-OCT-2000; 2000US-0240409P.  
XX PA (INCV-) INCYTE CORP.  
XX PI Furness LM, Buchbinder JL;  
XX DR WPI; 2004-068610/07.  
XX PT Combination useful for preparing a composition for treating liver  
XX disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,  
XX PT comprises cDNAs that are differentially expressed in response to steroid  
XX treatment.  
XX Claim 1; SEQ ID NO 1103; 141pp; English.

CC The invention relates to a combination comprising cDNAs that are  
CC differentially expressed in response to steroid treatment. Also included  
CC are the following: a high throughput method for using a cDNA to detect  
CC differential expression of nucleic acids in a sample; and a high  
CC throughput method of screening molecules or compounds to identify a  
CC ligand that specifically binds a cDNA. The sample is from a subject with  
CC Wilson disease and comparison of a standard defines a stage of that  
CC disease. The high throughput method of screening molecules or compounds  
CC to identify a ligand that specifically binds a cDNA comprises: combining  
CC the combination with molecules or compounds under conditions to allow  
CC specific binding; and detecting specific binding between each cDNA and at  
CC least one molecule or compound. The molecules or compounds are regulatory  
CC proteins. The combination is useful for preparing a composition for  
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis  
CC or hepatitis. The present sequence represents a human cDNA which is  
CC differentially expressed in steroid-induced C3A liver cells. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
XX at seqdata.uspto.gov/sequence.html.

XX Sequence 2797 BP; 866 A; 551 C; 595 G; 785 T; 0 U; 0 Other;  
Query Match 74.2%; Score 2315; DB 12; Length 2797;  
Best Local Similarity 99.8%; Pred No. 0;  
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

Qy 335 CCTGGAGTCAGCTTAAAGCTCTTGGCGATACCAAGAAATATCATGGCTACATGATGG 394  
Db 1 CCTGGAGTCAGCTTAAAGCTCTTGGCGATACCAAGAAATATCATGGCTACATGATGG 60  
Qy 395 CTAAGGCACCAATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 454  
Db 61 CTAAGGCACCAATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 120  
Qy 455 ACAGAAATCTATACCTTGGCCATGTCGTGTGAAGACAGAGAAATACACTGTTTTATCTCG 514  
Db 121 ACAGAAATCTATACCTTGGCCATGTCGTGTGAAGACAGAGAAATACACTGTTTTATCTCG 180  
Qy 515 AAATTCCTAAATCTCAATAGACGACAGTCTTAATGCTCTCTTGAAGCCCTCTTTTGG 574  
Db 181 AAATTCCTAAATCTCAATAGACGACAGTCTTAATGCTCTCTTGAAGCCCTCTTTTGG 240  
Qy 575 ATCTTTTTCAGGCAACACTGGACTATGGAATGTTATCTCGAAGAGGAACATATTAAAG 634  
Db 241 ATCTTTTTCAGGCAACACTGGACTATGGAATGTTATCTCGAAGAGGAACATATTAAAG 300  
Qy 635 AAAGAAAACGCTTGGAAACAGTCGGAATGCTCTTACGATTATCACCAGGAAGTGGAA 694  
Db 301 AAAGAAAACGCTTGGAAACAGTCGGAATGCTCTTACGATTATCACCAGGAAGTGGAA 360

Qy 695 CATTTCTGTTTCAGCCGCTAGTGAATTTATCACTGTAAGATCGAGGGCCACAAGGAT 754  
Db 361 CATTTCTGTTTCAGCCGCTAGTGAATTTATCACTGTAAGATCGAGGGCCACAAGGAT 420  
Qy 755 TTACGCAACACCTTTAAGGCCCAATCTAGTGAAGAACTAGTTGTCCCAACATACGAGATGG 814  
Db 421 TTACGCAACACCTTTAAGGCCCAATCTAGTGAAGAACTAGTTGTCCCAACATACGAGATGG 480  
Qy 815 ATCCAAAATTAATGCCCGCTGATCCAGACTGGAATGCTTTTATACATAGCAACGATATTT 874  
Db 481 ATCCAAAATTAATGCCCGCTGATCCAGACTGGAATGCTTTTATACATAGCAACGATATTT 540  
Qy 875 GGATATCTAACTGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTAG 934  
Db 541 GGATATCTAACTGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTAG 600  
Qy 935 CCACATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAAGAGAAT 994  
Db 601 CCACATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAAGAGAAT 660  
Qy 995 TTGATAGATATTCTGGCTATTGGTGTGTCAAAAAGCTGAAACAACTCCCAAGTGTGGTA 1054  
Db 661 TTGATAGATATTCTGGCTATTGGTGTGTCAAAAAGCTGAAACAACTCCCAAGTGTGGTA 720  
Qy 1055 AAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGTGGAATTTATTCATGTTA 1114  
Db 721 AAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGTGGAATTTATTCATGTTA 780  
Qy 1115 CATCCCTATGTTGGAAACAAGG-AGGCGACAGATTCATTCCTGTTATCTTAAACAGGTACA 1173  
Db 781 CATCCCTATGTTGGAAACAAGGCGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 840  
Qy 1174 GCAAACTCTAAAGTCACTTTAAGATGTCAGAAATGATGATGCTGGAAGAGATC 1233  
Db 841 GCAAACTCTAAAGTCACTTTAAGATGTCAGAAATGATGATGCTGGAAGAGATC 900  
Qy 1234 ATAGATGTCATAGATGAAGAACTAATTCACCTTTGAGATCTTATTTGAAGAGTTGAA 1293  
Db 901 ATAGATGTCATAGATGAAGAACTAATTCACCTTTGAGATCTTATTTGAAGAGTTGAA 960  
Qy 1294 TATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAAATA---TGCTTGGTCCATCTACT 1349  
Db 961 TATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAAATA---TGCTTGGTCCATCTACT 1020  
Qy 1350 AGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGT 1409  
Db 1021 AGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGT 1080  
Qy 1410 AGAAGATGATGTTATGGAAGGCGAGACTCATTTGAGTCACTGCTGATCTGTGACGCC 1469  
Db 1081 AGAAGATGATGTTATGGAAGGCGAGACTCATTTGAGTCACTGCTGATCTGTGACGCC 1140  
Qy 1470 ACTAAATATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCAATGT 1529  
Db 1141 ACTAAATATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCAATGT 1200  
Qy 1530 TTTTCCCAAAGTCAAGAGAGGAAATTTAGTATTTATTTTGGCTCTGAAATGCAAAACAGG 1589  
Db 1201 TTTTCCCAAAGTCAAGAGAGGAAATTTAGTATTTATTTTGGCTCTGAAATGCAAAACAGG 1260  
Qy 1590 TTTCCGTCAATTTATACAAATTTACATCATTTTAAAGGAAGCAATATATAAGATCCAG 1649  
Db 1261 TTTCCGTCAATTTATACAAATTTACATCATTTTAAAGGAAGCAATATATAAGATCCAG 1320  
Qy 1650 TGGTGGCTGCTCCAGTGTATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTTAC 1709  
Db 1321 TGGTGGCTGCTGCTCCAGTGTATTTCAAGTGTCTTATCAAGAGAGATAGCAATTTAC 1380  
Qy 1710 CAGTGTGATGGAGAGTTCTTGGCGGCAATGATCTTAATATCCAAATGATGAAAGTCAG 1769  
Db 1381 CAGTGTGATGGAGAGTTCTTGGCGGCAATGATCTTAATATCCAAATGATGAAAGTCAG 1440



QY 1770 AAGCTGCTATATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCACTGTACGTAGT 1829  
Db 1441 AAGCTGCTATATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCACTGTACGTAGT 1500  
QY 1830 CAGTTACGTAATCCTGGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTG 1889  
Db 1501 CAGTTACGTAATCCTGGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTG 1560  
QY 1890 CTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTATACCAAGAGAAATCCACA 1949  
Db 1561 CTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTATACCAAGAGAAATCCACA 1620  
QY 1950 CTGTGTGCTCCCTTTTCAAGCTATCAAGTCTGAAGTACCCAACTTGCAAAAACAAAGGA 2009  
Db 1621 CTGTGTGCTCCCTTTTCAAGCTATCAAGTCTGAAGTACCCAACTTGCAAAAACAAAGGA 1680  
QY 2010 ATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGAGCTATATCTCTCCAGAAAT 2069  
Db 1681 ATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGAGCTATATCTCTCCAGAAAT 1740  
QY 2070 TTTCTCTTTTGAAGTACTACTCGATTTTACATTTGATGGATGCTCTTACAAAGCCTCATGA 2129  
Db 1741 TTTCTCTTTTGAAGTACTACTCGATTTTACATTTGATGGATGCTCTTACAAAGCCTCATGA 1800  
QY 2130 TCTACAGCTGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGTGCA 2189  
Db 1801 TCTACAGCTGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGTGCA 1860  
QY 2190 GTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCT 2249  
Db 1861 GTTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCT 1920  
QY 2250 AdGTTATGCTGTTGCTAGTATAGACAAACAGGGGATCCTGTACACGAGGGCTTAAATTTGA 2309  
Db 1921 AGTTATGCTGTTGCTAGTATAGACAAACAGGGGATCCTGTACACGAGGGCTTAAATTTGA 1980  
QY 2310 AdGGCCCTTTAAATATATAATATGAGTCAAAATAGAAATTTGACATCAGGTGGAAGACTCCA 2369  
Db 1981 AGGGCCCTTTAAATATATAATATGAGTCAAAATAGAAATTTGACATCAGGTGGAAGACTCCA 2040  
QY 2370 ATATCTAGCTTCTCGATATGATTTTCAATTTGATTTAGTCTGTTGGGCATCCAGGCTGGTC 2429  
Db 2041 ATATCTAGCTTCTCGATATGATTTTCAATTTGATTTAGTCTGTTGGGCATCCAGGCTGGTC 2100  
QY 2430 CTATGGAGGATACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTCAGGGTTGC 2489  
Db 2101 CTATGGAGGATACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTCAGGGTTGC 2160  
QY 2490 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTA 2549  
Db 2161 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTA 2220  
QY 2550 TATGGGTCACTGACCAAGATGAA CAGGGCTATTTAGTATGATCTGTGGCCATGCAAGC 2609  
Db 2221 TATGGGTCACTGACCAAGATGAA CAGGGCTATTTAGTATGATCTGTGGCCATGCAAGC 2280  
QY 2610 AGAAAGTTCCTCTGACCAATCTTTTACTGCTCTTACATGTTTCTCTGGATGAGAA 2669  
Db 2281 AGAAAGTTCCTCTGACCAATCTTTTACTGCTCTTACATGTTTCTCTGGATGAGAA 2340  
QY 2670 TGTTCATTTTGCATACCAATACAGTATTTACTGAGTTTTTTTGTGAGGGCTGGAAGCCATA 2729  
Db 2341 TGTTCATTTTGCATACCAATACAGTATTTACTGAGTTTTTTTGTGAGGGCTGGAAGCCATA 2400  
QY 2730 TGAATTACAGATCTATCTCAGGAGACACACAGTAAAGATTCCTGATTCGGAGAACCA 2789  
Db 2401 TGAATTACAGATCTATCTCAGGAGACACACAGTAAAGATTCCTGATTCGGAGAACCA 2460  
QY 2790 TTATGAACTGCTATTTTGGCACTTCTCAAGAAACCTTGGATCAGCTATTTGCTGCTCT 2849  
Db 2461 TTATGAACTGCTATTTTGGCACTTCTCAAGAAACCTTGGATCAGCTATTTGCTGCTCT 2520  
QY 2850 AAAAGTGATATAATTTTGA CCGTGTGTAGAACTCTCTGGTATATACACTGGCTATTTAAACCA 2909

Db 2521 AAAAGTGATATAATTTTGA CCGTGTGTAGAACTCTCTGGTATATACACTGGCTATTTAAACCA 2580  
QY 2910 ATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCATCTTTTGATACCTGCCAT 2969  
Db 2581 ATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCATCTTTTGATACCTGCCAT 2640  
QY 2970 GTAACTACTCTCTCTGAAAATAAATGTGTGCCATGTCAGGGGCTCTACGGTTTGTGGTAGT 3029  
Db 2641 GTAACTACTCTCTCTGAAAATAAATGTGTGCCATGTCAGGGGCTCTACGGTTTGTGGTAGT 2700  
QY 3030 AATCTAATACCTTTAAACCCACATGCTCAAAATCAATCAATGATATATCTCTGAGAGACCA 3089  
Db 2701 AATCTAATACCTTTAAACCCACATGCTCAAAATCAATGATATATCTCTGAGAGACCA 2760  
QY 3090 GCAATACCATTAAGAAATTTACTAAAAA 3120  
Db 2761 GCAATACCATTAAGAAATTTACTAAAAA 2791  
RESULT 11  
ACA92425  
ID ACA92425 standard; DNA; 2929 BP.  
XX  
AC ACA92425;  
DT  
DT 15-JUL-2003 (first entry)  
XX  
DE DNA encoding human PMM-10.  
XX  
KW Human; protein modification and maintenance molecule; PMM; cancer;  
cell proliferation disorder; atherosclerosis; neurological disorder;  
epilepsy; Huntington's disease; stroke; immune disorder; allergy;  
inflammatory disorder; AIDS; developmental disorder; hypothyroidism;  
Cushing's syndrome; gastrointestinal disorder; epithelial disorder;  
infection; cystostatic; antiatherosclerotic; anticonvulsant; nootropic;  
neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnery;  
antiinflammatory; thymimetic; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2003031939-A2.  
XX  
PD 17-APR-2003.  
XX  
PF 11-OCT-2002; 2002WO-US032850.  
XX  
PR 12-OCT-2001; 2001US-0329689P.  
PR 25-OCT-2001; 2001US-0335703P.  
PR 09-NOV-2001; 2001US-034887P.  
PR 28-NOV-2001; 2001US-0334145P.  
PR 06-DEC-2001; 2001US-0337451P.  
PR 14-DEC-2001; 2001US-0340584P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
Tran UK, Becha SD, Duggan BM, Lee BA, Griffin JA, Li JX;  
Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;  
Pi Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;  
Pi Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
XX  
DR WPI; 2003-430274/40.  
DR P-PSDB; ABU92030.  
XX  
PT New human protein modification and maintenance molecules (PMM), useful  
for diagnosing; treating and preventing diseases or conditions associated  
PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or  
infections.  
XX  
PS Claim 5; Page 289; 311pp; English.  
XX  
CC The present invention relates to the isolation of human protein



Db	2374	CCCTGACCAAGATGAACAGGCGCTATTCTTAGGATCTGTGGCCATGCAAGACAGAAAAGTT	2433
Qy	2619	CCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGCCATT	2678
Db	2434	CCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGCCATT	2493
Qy	2679	TGCACATACCACTATATTACTGAGTTTTTTAGTGAAGGCTGGAAGCCATATGATTTACA	2738
Db	2494	TGCACATACCACTATATTACTGAGTTTTTTAGTGAAGGCTGGAAGCCATATGATTTACA	2553
Qy	2739	GATCTATCTCCAGGAGACACAGCATAAAGATTTCTGAAATCGGGAGAACATATGAAT	2798
Db	2554	GATCTATCTCCAGGAGACACAGCATAAAGATTTCTGAAATCGGGAGAACATATGAAT	2613
Qy	2799	GGATCTTTTGCACTACCTTCAAGAAAACCTTTGGATCAGGATATTCCTGCTCTAAAGTGAT	2858
Db	2614	GCATCTTTTGCACTACCTTCAAGAAAACCTTTGGATCAGGATATTCCTGCTCTAAAGTGAT	2673
Qy	2859	ATAATTTTGACCTGTAGAACTCTCTGATATACACTGGCTATTAAACCAATGAGGAG	2918
Db	2674	ATAATTTTGACCTGTAGAACTCTCTGATATACACTGGCTATTAAACCAATGAGGAG	2733
Qy	2919	TTTAATCAACAGAAAACACAGAAATTTGATCATCAATTTTGATACCTGCCATGTAACAT	2978
Db	2734	TTTAATCAACAGAAAACACAGAAATTTGATCATCAATTTTGATACCTGCCATGTAACAT	2793
Qy	2979	ATCTCTGAAATTAATTTGGTGGCCATGAGGGTCTACGGTTTGTGGTAGTAATCTAATA	3038
Db	2794	ATCTCTGAAATTAATTTGGTGGCCATGAGGGTCTACGGTTTGTGGTAGTAATCTAATA	2853
Qy	3039	CTTTAACCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCA	3098
Db	2854	CTTTAACCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCA	2913
Qy	3099	TAAGAAATTAATAAAAA	3114
Db	2914	TAAGAAATTAATAAAAA	2929
RESULT 12			
ACA92421			
ID	ACA92421	standard; DNA; 2952 BP.	
XX	ACA92421;		
AC	ACA92421;		
XX	15-JUL-2003	(first entry)	
DT	15-JUL-2003	(first entry)	
XX	DNA encoding human PMM-6.		
DE	Human; protein modification and maintenance molecule; PMM; cancer;		
XX	cell proliferation disorder; atherosclerosis; neurological disorder;		
KW	epilepsy; Huntington's disease; stroke; immune disorder; allergy;		
KW	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;		
KW	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;		
KW	infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;		
KW	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;		
KW	antiinflammatory; thymimetic; gene; ds.		
XX	Homo sapiens.		
OS	WO2003031939-A2.		
XX	17-APR-2003.		
PN	11-OCT-2003; 2002WO-US032850.		
XX	12-OCT-2001; 2001US-0329689P.		
PR	25-OCT-2001; 2001US-0335703P.		
PR	09-NOV-2001; 2001US-0348887P.		
PR	28-NOV-2001; 2001US-0334145P.		
PR	06-DEC-2001; 2001US-0337451P.		
PR	14-DEC-2001; 2001US-0340584P.		

PA	(INCY-) INCYTE GENOMICS INC.				
XX	Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;				
PI	Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;				
PI	Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;				
PI	Marquis JP, Rao MG, Richardson TW, Tang TY, Jin P, Chien D;				
PI	Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;				
XX	WPI; 2003-430274/40.				
DR	P-PSDB; ABU92026.				
XX	New human protein modification and maintenance molecules (PMM), useful				
PT	for diagnosing, treating and preventing diseases or conditions associated				
PT	with the aberrant PMM expression e.g. cancer, atherosclerosis, or				
PT	infections.				
PS	Claim 5; Page 285-286; 311pp; English!				
XX	The present invention relates to the isolation of human protein				
CC	modification and maintenance molecules (PMM), and the polynucleotide				
CC	sequences encoding them. A total of 40 PMM polypeptides (designated PMM				
CC	-1 to PMM-40) are disclosed. The sequences of the invention are useful				
CC	for diagnosing a condition or disease associated with the expression of				
CC	PMM in a subject, preparing a polyclonal or monoclonal antibody, and				
CC	generating an expression profile of a sample containing the				
CC	polynucleotides. The diseases or conditions associated with decreased				
CC	expression or overexpression of PMM are cell proliferation disorders				
CC	(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,				
CC	Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,				
CC	allergies), developmental disorders (e.g. hypothyroidism, Cushing's				
CC	syndrome), gastrointestinal or epithelial disorders, and infections. The				
CC	PMM polypeptides or their fragments are useful in screening compounds				
CC	for effectiveness as agonists or antagonists of the polypeptides, or in				
CC	altering the expression of the target polynucleotide and compounds that				
CC	specifically bind to, or modulate the activity of the polypeptide.				
CC	ACA92416-ACA92455 encode the human PMM polypeptides of the invention				
XX	Sequence 2952 BP; 886 A; 604 C; 664 G; 798 T; 0 U; 0 Other;				
Query Match	72.8%; Score 2271; DB 10; Length 2952;				
Best Local Similarity	99.9%; Pred. No! 0;				
Matches 2321; Conservative	0; Mismatches 1; Indels 0; Gaps 0;				
Qy	11 GCCTCCGAGCCCAAGCCCGCTCTACTGCGCGCGTGTCTTCTAGTCCGCGTTCGCCGC	70			
Db	1 GCCTCCGAGCCCAAGCCCGCTCTACTGCGCGCGTGTCTTCTAGTCCGCGTTCGCCGC	60			
Qy	71 CTGGTTTGTACCG	130			
Db	61 CTGGTTTGTACCG	120			
Qy	131 GGCGCAGCATGAAGCG	190			
Db	121 GGCGCAGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	180			
Qy	191 GCCCGGGGGAAGAAATGCAATGCGAGCGAATGGAACACAGACAGCGTGGTCTTG	250			
Db	181 GCCCGGGGGAAGAAATGCAATGCGAGCGAATGGAACACAGACAGCGTGGTCTTG	240			
Qy	251 AGATATTTGAACTGCGGCGCTGTGAGGAGAAATTTGAATACAGGATCGGCTAAATGG	310			
Db	241 AGATATTTGAACTGCGGCGCTGTGAGGAGAAATTTGAATACAGGATCGGCTAAATGG	300			
Qy	311 AGCCTTTTATTTAGTGTAGCGGTATTTCTCTGGAGTCAGCTTAAAAAGTCTTCCCATACCA	370			
Db	301 AGCCTTTTATTTAGTGTAGCGGTATTTCTCTGGAGTCAGCTTAAAAAGTCTTCCCATACCA	360			
Qy	371 GAAATATCATGCTACATGATGCTAAGCCACCATGATTTTCTGTTGTGAGAGGA	430			
Db	361 GAAATATCATGCTACATGATGCTAAGCCACCATGATTTTCTGTTGTGAGAGGA	420			
Qy	431 ATGATCCAGATGACCTTATTTCAGACAGATCTATTACCTTCCCATGCTCTGTGTGAGAACA	490			

Db 421 ATGATCCAGATGGACCTCAITTCAGACAGAATCTATTACCTTGCATGTCGTGGAGACA 480  
QY 491 GAGAAATACACTGTTTTATTTCTGAATTCCTCAAAATCCCAAACTATCAATAGACAGCAGCTCTAA 550  
Db 481 GAGAAATACACTGTTTTATTTCTGAATTCCTCAAAATCCCAAACTATCAATAGACAGCAGCTCTAA 540  
QY 551 TGCTCTCTTTGGAAGCCCTCTTTTGGATCTTTTTCAGGACACACTGGACATCTGGAATGATT 610  
Db 541 TGCTCTCTTTGGAAGCCCTCTTTTGGATCTTTTTCAGGACACACTGGACATCTGGAATGATT 600  
QY 611 CTCGAGAGAGACACTAATTAAGAGAAAGAAACCGCATTTGGAACAGCTCGGAATTCCTTCTT 670  
Db 601 CTCGAGAGAGACACTAATTAAGAGAAAGAAACCGCATTTGGAACAGCTCGGAATTCCTTCTT 660  
QY 671 ACGATATACCAAGAGAGTGGACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACG 730  
Db 661 ACGATATACCAAGAGAGTGGACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACG 720  
QY 731 TAAAGATGGAGGGCCCAAGAGATTTACGAACAACTTTAAGGCCCAATCTAGTGGAAA 790  
Db 721 TAAAGATGGAGGGCCCAAGAGATTTACGAACAACTTTAAGGCCCAATCTAGTGGAAA 780  
QY 791 CTAGTTGTCGCAACATACGATGGATTCGCAAAATTTATGCCCGCTGATCCAGACTGGATTG 850  
Db 781 CTAGTTGTCGCAACATACGATGGATTCGCAAAATTTATGCCCGCTGATCCAGACTGGATTG 840  
QY 851 CTTTATATACATAGCAACGATTTTGGATATCTAAATCATCGTAACCAAGAGAAAGAGAC 910  
Db 841 CTTTATATACATAGCAACGATTTTGGATATCTAAATCATCGTAACCAAGAGAAAGAGAC 900  
QY 911 TCATTTATGTCACAAATAGCTAGCCCAATGGAAGAGATGCGCAGATCAGCTGGAGTCG 970  
Db 901 TCATTTATGTCACAAATAGCTAGCCCAATGGAAGAGATGCGCAGATCAGCTGGAGTCG 960  
QY 971 CTACCTTTGTTCTCCAAAGAAATTTGATAGATATCTGGCTATTTGGTGTGTCGCAAAAG 1030  
Db 961 CTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTGGCTATTTGGTGTGTCGCAAAAG 1020  
QY 1031 CTGAACAACTCCAGTGTGTGTAATTTCTAGATATCTATATGAGAAATGATGAT 1090  
Db 1021 CTGAACAACTCCAGTGTGTGTAATTTCTAGATATCTATATGAGAAATGATGAT 1080  
QY 1091 CTGAGTGGAAATTTATTCATGTTATCATCCCTATTTGGAAACAGAGAGGCGAGTTCA 1150  
Db 1081 CTGAGTGGAAATTTATTCATGTTATCATCCCTATTTGGAAACAGAGAGGCGAGTTCA 1140  
QY 1151 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATA 1210  
Db 1141 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATA 1200  
QY 1211 TGAATGATCTGAAGAGAGATCATAGATGTCATAGATAAGGAACTAATTCACCTTTTG 1270  
Db 1201 TGAATGATCTGAAGAGAGATCATAGATGTCATAGATAAGGAACTAATTCACCTTTTG 1260  
QY 1271 AGATTCATTTTGAAGAGTTGAATATATTCAGAGCTGGATGGACTCCTGAGGGAAAT 1330  
Db 1261 AGATTCATTTTGAAGAGTTGAATATATATTCAGAGCTGGATGGACTCCTGAGGGAAAT 1320  
QY 1331 ATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCAC 1390  
Db 1321 ATGCTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCAC 1380  
QY 1391 CTGAATTTATTTTCCAGTGAAGATGATGTTATGGAAGGAGAGACTCAATTTAGTCTCAG 1450  
Db 1381 CTGAATTTATTTTCCAGTGAAGATGATGTTATGGAAGGAGAGACTCAATTTAGTCTCAG 1440  
QY 1451 TGCTCTGATTTCTGAGCGCCTAATTTATCTATGAAGAAACACAGACATCTGGATAATA 1510  
Db 1441 TGCTCTGATTTCTGAGCGCCTAATTTATCTATGAAGAAACACAGACATCTGGATAATA 1500  
QY 1511 TCCATGACATCTTTTCATGTTTTTCCCAAGTCCCAAGAGGAAATGAGTTATTTTGG 1570  
Db 1501 TCCATGACATCTTTTCATGTTTTTCCCAAGTCCCAAGAGGAAATGAGTTATTTTGG 1560

QY 1571 CCTCTGAATGCAAAAACAGGTTTCCGTCAATTTACAAAATTACATCTATTTTAAAGGAAA 1630  
Db 1561 CCTCTGAATGCAAAAACAGGTTTCCGTCAATTTACAAAATTACATCTATTTTAAAGGAAA 1620  
QY 1631 GCAAAATATAAACGATCCAGTGGTGGCTGCTCCCTGCTCCAAAGTGAATTTCAAGTGTCTATCA 1690  
Db 1621 GCAAAATATAAACGATCCAGTGGTGGCTGCTCCCTGCTCCAAAGTGAATTTCAAGTGTCTATCA 1680  
QY 1691 AAGAGGAGATAGCAAAATTTACCAAGTGGTGAATGGAAAGTTCTTGGCCGGCATCGATCTAATA 1750  
Db 1681 AAGAGGAGATAGCAAAATTTACCAAGTGGTGAATGGAAAGTTCTTGGCCGGCATCGATCTAATA 1740  
QY 1751 TCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCCAAAAGACTCCCTTTAG 1810  
Db 1741 TCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCCAAAAGACTCCCTTTAG 1800  
QY 1811 AGCATCACTGTACGTAGTCAGTTACGTAAATCTTGGAGAGGTGACAGGGCTGACTGACC 1870  
Db 1801 AGCATCACTGTACGTAGTCAGTTACGTAAATCTTGGAGAGGTGACAGGGCTGACTGACC 1860  
QY 1871 GTGGCTACTCACATTTCTGCTGCATCAGTCAGTCTGACTTCTTTTAAAGTAAATATA 1930  
Db 1861 GTGGCTACTCACATTTCTGCTGCATCAGTCAGTCTGACTTCTTTTAAAGTAAATATA 1920  
QY 1931 GTAACCCAGAGATCCCACTGTGCTCCCTTTACAGCTATCAAGTCTCGAAGATGACC 1990  
Db 1921 GTAACCCAGAGATCCCACTGTGCTCCCTTTACAGCTATCAAGTCTCGAAGATGACC 1980  
QY 1991 CAATTTGCAAAAACAAAGGAATTTTGGGCCCAATTTTGGATTCAGAGGTGACAGGGCTGACTGACC 2050  
Db 1981 CAATTTGCAAAAACAAAGGAATTTTGGGCCCAATTTTGGATTCAGAGGTGACAGGGCTGACTGACC 2040  
QY 2051 ACTATATCTCCTCAGAAATTTTCTCTTTTGAAGTACTTCTGATTTTACATTTGATGGGA 2110  
Db 2041 ACTATATCTCCTCAGAAATTTTCTCTTTTGAAGTACTTCTGATTTTACATTTGATGGGA 2100  
QY 2111 TGCTCTACAGCTCATGATCTACAGCTGGAAGAAATATCCCTACTGCTGCTTCATAT 2170  
Db 2101 TGCTCTACAGCTCATGATCTACAGCTGGAAGAAATATCCCTACTGCTGCTTCATAT 2160  
QY 2171 ATGGTGGTCTCCTCAGGTGAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCGCT 2230  
Db 2161 ATGGTGGTCTCCTCAGGTGAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCGCT 2220  
QY 2231 TGAATACCTTAGCCTCTCTAGGTTATGTTGGTTAGTAGAGACACAGGGGATCTGTC 2290  
Db 2221 TGAATACCTTAGCCTCTCTAGGTTATGTTGGTTAGTAGAGACACAGGGGATCTGTC 2280  
QY 2291 ACCGAGGGCTTAAATTTTGAAGGCCCTTTAAATATAAAATGG 2332  
Db 2281 ACCGAGGGCTTAAATTTTGAAGGCCCTTTAAATATAAAATGG 2322

## RESULT 13

AD116386

ID AD116386 standard; DNA; 2696 BP.

XX AC AD116386;

XX DT 22-APR-2004 (first entry)

XX DE Human protein modification and maintenance molecule (PMM) gene #19.

XX KW human; protein modification and maintenance molecule; PMM;

XX KW gastrointestinal disorders; peptic ulcer; Crohn's disease;

XX KW cardiovascular disorders; hypertension; congenital heart disease;

XX KW autoimmune disease; inflammatory disease; AIDS; anaemia;

XX KW developmental disorder; Cushing's syndrome; tubular acidosis;

XX KW epithelial disorder; eczema; scabies; neurological disorder;

XX KW Alzheimer's disease; multiple sclerosis; infection; cancer; ds.

XX OS Homo sapiens.



QY 1513 CATGACATCTTTTCATGTTTTTCCCAAGTCACGAGAGGAATTTGAGTTATTTTGGC 1572  
Db |||||  
QY 1502 CATGACATCTTTTCATGTTTTTCCCAAGTCACGAGAGGAATTTGAGTTATTTTGGC 1561  
Db |||||  
QY 1573 TCTGAATGCAAAACAGAGTTTCCTGCTATTTATACAAATTTACATTTTAAAGGAAGC 1632  
Db |||||  
QY 1562 TCTGAATGCAAAACAGAGTTTCCTGCTATTTATACAAATTTACATTTTAAAGGAAGC 1621  
Db |||||  
QY 1633 AAATATAACGATCCAGTGTGGGCTGCTGCTCCAAAGTATTTCAAAGTGTCCATCAAA 1692  
Db |||||  
QY 1622 AAATATAACGATCCAGTGTGGGCTGCTGCTCCAAAGTATTTCAAAGTGTCCATCAAA 1681  
Db |||||  
QY 1693 GAGGAGATAGCAATTTACATGTGTGAATGGAGTTCTTGGCCGATGATTAATATC 1752  
Db |||||  
QY 1682 GAGGAGATAGCAATTTACATGTGTGAATGGAGTTCTTGGCCGATGATTAATATC 1741  
Db |||||  
QY 1753 CAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAG 1812  
Db |||||  
QY 1742 CAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAGAG 1801  
Db |||||  
QY 1813 CATCACTGTACGTTAGTTCAGTTACGTTAAATCTCTGGAGAGGTGACAAAGGCTGACGCGT 1872  
Db |||||  
QY 1802 CATCACCTGTACGTTAGTTCAGTTACGTTAAATCTCTGGAGAGGTGACAAAGGCTGACGCGT 1861  
Db |||||  
QY 1873 GGCTACTCACATTTCTGCTCATCAGTCAGTCTGACTCTCTTTTATAAGTAACTATAGT 1932  
Db |||||  
QY 1862 GGCTACTCACATTTCTGCTCATCAGTCAGTCTGACTCTCTTTTATAAGTAACTATAGT 1921  
Db |||||  
QY 1933 AACCAAGAGATCCACATGTGTGCTCTTTTACAGCTATCAAGTCTCTGAGATGACCCA 1992  
Db |||||  
QY 1922 AACCAAGAGATCCACATGTGTGCTCTTTTACAGCTATCAAGTCTCTGAGATGACCCA 1981  
Db |||||  
QY 1993 ACTTGCAAAACAAAGGATTTTGGCCACCAATTTGGATTCAGAGGCTCTCTCTGAC 2052  
Db |||||  
QY 1982 ACTTGCAAAACAAAGGATTTTGGCCACCAATTTGGATTCAGAGGCTCTCTCTGAC 2041  
Db |||||  
QY 2053 TATACTCTCCAGAAATTTCTTTTGAAGTACTACTGGAATTAATTGATGAGATG 2112  
Db |||||  
QY 2042 TATACTCTCCAGAAATTTCTTTTGAAGTACTACTGGAATTAATTGATGAGATG 2101  
Db |||||  
QY 2113 CTCTACAGCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTTCAATAT 2172  
Db |||||  
QY 2102 CTCTACAGCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTTCAATAT 2161  
Db |||||  
QY 2173 GGTGGTCTCAGTGAGTGTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTG 2232  
Db |||||  
QY 2162 GGTGGTCTCAGTGAGTGTGGTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTG 2221  
Db |||||  
QY 2233 AAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCTGTAC 2292  
Db |||||  
QY 2222 AAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCTGTAC 2281  
Db |||||  
QY 2293 CGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATG 2331  
Db |||||  
QY 2282 CGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATG 2320  
Db |||||

RESULT 14  
ABK83325  
ID ABK83325 standard; cDNA; 4523 BP.  
XX  
AC ABK83325;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE cDNA encoding human DPRP-1 splice variant #1.  
XX  
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;  
KW DPRP; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;  
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;  
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;  
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;

KW dyskinesia; reproductive disorder; inflammatory disorder;  
XX metabolic disorder; gene; ss.  
OS Homo sapiens.

XX WO200231134-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US031874.

XX 12-OCT-2000; 2000US-0240117P.

XX (FERR ) FERRING BV.

XX Qi S, Akinsanya KO, Riviere PJ, Junien J;

XX WPI; 2002-444178/47.

XX P-PSDB; ABG61594.

XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding  
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and  
PT viral infections, cancers, allergies, neurological disorders, or pain.

XX Disclosure; Page 61-62; 113pp; English.

XX The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins  
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic  
CC acids encoding them are useful for treating infections such as fungal,  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraines, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322  
CC -ABK83343 encode human DPRP proteins

XX Sequence 4523 BP; 1384 A; 828 C; 940 G; 1371 T; 0 U; 0 Other;

Query Match 68.2%; Score 2128; DB 6; Length 4523;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGTGTCTAAAGCTCCGAGGCCCAAGGCCGCTGCTACTGCGCCGCTCTCTTAGTGCCG 60  
Db |||||

QY 1 AAGTGTCTAAAGCTCCGAGGCCCAAGGCCGCTGCTACTGCGCCGCTCTCTTAGTGCCG 60  
Db |||||

QY 61 CGTTCCCGCCCTGGGTTGTACCCGCGCCGCGGAGGAGCCACTGCAACGAGNCCG 120  
Db |||||

QY 61 CGTTCCCGCCCTGGGTTGTACCCGCGCCGCGGAGGAGCCACTGCAACGAGNCCG 120  
Db |||||

QY 121 GAGTGGAGGGGGCGGAGCATGAGGGGCGAGGCCGCTCATAGCGCAGTGGGACCG 180  
Db |||||

QY 121 GAGTGGAGGGGGCGGAGCATGAGGGGCGAGGCCGCTCATAGCGCAGTGGGACCG 180  
Db |||||

QY 181 TCCGGCGGGGGCGGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACACAGAACAG 240  
Db |||||

QY 181 TCCGGCGGGGGCGGGGGAGGAAATGCAACATGGCAGCAGCAATGGAACACAGAACAG 240  
Db |||||

QY 241 CTGGGGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300  
Db |||||

QY 241 CTGGGGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300  
Db |||||

QY 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTTCTTGGAGTCAAGCTTAAAAAGCTGCTT 360  
Db |||||

QY 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTTCTTGGAGTCAAGCTTAAAAAGCTGCTT 360  
Db |||||

QY 361 GCGGATACAGAAAAATATCATGGGTACATGATGGCTAAGGCACCATATGATTTTCATGTTT 420  
Db |||||



```
Db 361 GCCATACCAGAAAATATCATGGCTACATGATGGCTAAGGCCACCACCATGATTTTCATGTTT 420
QY 421 GTGAAGAGGAATGATCCAGATGACACCTCATTCAGACAGAACTATTATACCTTTGCCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGACACCTCATTCAGACAGAACTATTATACCTTTGCCATGCT 480
QY 481 GGTGAGAACAGAGAAAATACACTGTTTATTCTGAAATTCGCCAAATTCATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAAATACACTGTTTATTCTGAAATTCGCCAAATTCATCAATAGAGCA 540
QY 541 GCACTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCACTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
QY 601 GGAATGATTCTTCGAGAAGAGAACTATTAAAGAGAAAGAAAACGCAATTTGGAACAGTCGGA 660
Db 601 GGAATGATTCTTCGAGAAGAGAACTATTAAAGAGAAAGAAAACGCAATTTGGAACAGTCGGA 660
QY 661 ATTGCTTCTTACGATTATACCAAGGAAGTGGAACTATTTCTGTTTCAAGCCGGTAGTGA 720
Db 661 ATTGCTTCTTACGATTATACCAAGGAAGTGGAACTATTTCTGTTTCAAGCCGGTAGTGA 720
QY 721 ATTATATCACTGTAAGAGTGGAGGCCACAGGATTTACGCAACACCTTTTAAGGCCAAT 780
Db 721 ATTATATCACTGTAAGAGTGGAGGCCACAGGATTTACGCAACACCTTTTAAGGCCAAT 780
QY 781 CTAGTGGAAACTAGTGTCTCCAAACATACGATGGATTCGCAAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAAACTAGTGTCTCCAAACATACGATGGATTCGCAAAATTTATGCCCGCTGATCCA 840
QY 841 GACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACAGAGAA 900
Db 841 GACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTAAACATCGTAACAGAGAA 900
QY 901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCACAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCA 960
QY 961 GCTGGAGTCGTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCGTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
QY 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGTGTAATAATTTCTAGAAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGTGTAATAATTTCTAGAAATTTCTATATGAAGAA 1080
QY 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATCTGTGGAACCAAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATCTGTGGAACCAAGGAGG 1140
QY 1141 GAGATTCATTTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCATTTTAAAGATG 1200
Db 1141 GAGATTCATTTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCATTTTAAAGATG 1200
QY 1201 TCAGAAATPAATGATGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACATAAT 1260
Db 1201 TCAGAAATPAATGATGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACATAAT 1260
QY 1261 CAACCTTTTGAGATCTTATTGGAAGGAGTTGAAATATATTGCCAGAGCTGGATGCACTCT 1320
Db 1261 CAACCTTTTGAGATCTTATTGGAAGGAGTTGAAATATATTGCCAGAGCTGGATGCACTCT 1320
QY 1321 GAGGGAATATGCTGCTGCTCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Db 1321 GAGGGAATATATGCTTGGTTCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
QY 1381 TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATTTTGAAGAGGCAGAGACTC 1440
Db 1381 TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATTTTGAAGAGGCAGAGACTC 1440
QY 1441 ATTGAGTCAGTCCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500
Db 1441 ATTGAGTCAGTCCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500
```

```
QY 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAAGAGAAATTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAAGAGAAATTGAG 1560
QY 1561 TTTATTTTTCCTCTCGAATGCAAAAACAGGTTTCCGTCATTTTATATACAAATATACATCTATT 1620
Db 1561 TTTATTTTTCCTCTCGAATGCAAAAACAGGTTTCCGTCATTTTATATACAAATATACATCTATT 1620
QY 1621 TTAAGGAAAGCAAAATATAAAACGATCCAGTGGTGGGCTGCCCTGCCAAGTGATTTCAAG 1680
Db 1621 TTAAGGAAAGCAAAATATAAAACGATCCAGTGGTGGGCTGCCCTGCCAAGTGATTTCAAG 1680
QY 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
Db 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
QY 1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db 1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
QY 1801 TCCCTTTTAGAGCATCACTGTAGTCAGTTACGTAAATTCCTGGAGAGGTGAAGAAG 1860
Db 1801 TCCCTTTTAGAGCATCACTGTAGTCAGTTACGTAAATTCCTGGAGAGGTGAAGAAG 1860
QY 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACAATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTATA 1920
QY 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCCT 1980
Db 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCCT 1980
QY 1981 GAAGATGACCCCACTTCGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCACAGGT 2040
Db 1981 GAAGATGACCCCACTTCGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCACAGGT 2040
QY 2041 CCTCTTCTGACTATATCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTACA 2100
Db 2041 CCTCTTCTGACTATATCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTACA 2100
QY 2101 TTGTATGGGATGCTCTACAAGCCCTCATGATCTACAGCCCTGGAAAGAAATATCCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTACAAGCCCTCATGATCTACAGCCCTGGAAAGAAATATCCTACTGTG 2160
QY 2161 CTGTTTCATATATGTTGTC 2179
Db 2161 CTGTTTCATATATGTTGTC 2179
```

## RESULT 15

ABK83331

ID ABK83331 standard; cDNA; 4676 BP.

XX AC ABK83331;

XX DT 12-AUG-2002 (first entry)

XX CDNA encoding human DPRP-1 splice variant #7.

Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP; DPP4V; infection; human immunodeficiency virus; HIV-1; HIV-2; pain; diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke; heart failure; hypertension; urinary retention; osteoporosis; cancer; ulcer; allergy; cancer; psychotic disorder; neurological disorder; dyskinesia; reproductive disorder; inflammatory disorder; metabolic disorder; gene; ss.

OS Homo sapiens.

XX WO200231134-A2.

XX PD 18-APR-2002.



XX 12-OCT-2001; 2001WO-US031874.  
PF 12-OCT-2000; 2000US-0240117P.  
PR (FERR ) FERRING BV.  
XX  
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;  
XX WPI: 2002-444178/47.  
XX P-P8DB; ABG61600.  
DR  
XX  
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding  
PI the proteins, useful for treating e.g. fungal, bacterial, protozoan and  
PT viral infections, cancers, allergies, neurological disorders, or pain.  
XX  
XX Disclosure; Page 72-73; 113pp; English.  
XX  
XX The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins  
CC (DPPR). The dipeptidyl peptidase IV-related proteins (DPPR) and nucleic  
CC acids encoding them are useful for treating infections such as fungal,  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322  
CC -ABK83343 encode human DPPR proteins  
XX  
XX Sequence 4676 BP; 1424 A; 859 C; 979 G; 1414 T; 0 U; 0 Other;

Query Match 68.28; Score 2128; DB 6; Length 4676;  
Best Local Similarity 99.98; Pred. No. 0;  
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AAGTCTTAAGCCTCCAGGCCAAGCCGCTGACTGCGCGCGTCTTTAGTCCG 60  
DB 1 AAGTCTTAAGCCTCCAGGCCAAGCCGCTGACTGCGCGCGTCTTTAGTCCG 60  
QY 61 CGTTCCGCGCGCTGGGTTGTCCAGCGCGCGCGCGAGGAGCCACTGCAACAGGACCG 120  
DB 61 CGTTCCGCGCGCTGGGTTGTCCAGCGCGCGCGCGAGGAGCCACTGCAACAGGACCG 120  
QY 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGTCGGGACGG 180  
DB 121 GAGTGGAGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGTCGGGACGG 180  
QY 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAACACAGAACAG 240  
DB 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAACACAGAACAG 240  
QY 241 CTGGGTGTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAATCAGGATCGG 300  
DB 241 CTGGGTGTGAGATATTTGAACTCGGACTGTGAGGAGATATTTGAATCAGGATCGG 300  
QY 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGT 360  
DB 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGAGCTGT 360  
QY 361 GCCGATACAGAAATATCATGCTACATGATGGCTTAGGCCACCATGATTTTCATGTT 420  
DB 361 GCCGATACAGAAATATCATGCTACATGATGGCTTAGGCCACCATGATTTTCATGTT 420  
QY 421 GTGAGAGGATCATCCAGATGACCTCATTCAGACAGAACTATTACCTTGCATGCT 480  
DB 421 GTGAGAGGATCATCCAGATGACCTCATTCAGACAGAACTATTACCTTGCATGCT 480  
QY 481 GGTGAGACAGAAATATCATGTTTATTTCTGAAATCCCAAACTATCATAGAGCA 540  
DB 481 GGTGAGACAGAAATATCATGTTTATTTCTGAAATCCCAAACTATCATATCTATT 1620

DB 481 GGTGAGACAGAAATATCATGTTTATTTCTGAAATCCCAAACTATCATATAGACA 540  
QY 541 GCAGTCTTAATGCTCTCTCGAAGCCTCTTTTGGATCTTTTTCAGCGAACACTGACTAT 600  
DB 541 GCAGTCTTAATGCTCTCTCGAAGCCTCTTTTGGATCTTTTTCAGCGAACACTGACTAT 600  
QY 601 GGAATGTATTTCTCGAAGAAAGAACTATTAAAGAAAGAAAAACGATTTGGAACTCGGA 660  
DB 601 GGAATGTATTTCTCGAAGAAAGAACTATTAAAGAAAGAAAAACGATTTGGAACTCGGA 660  
QY 661 ATTGCTTTCTTACGATTTATCAACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA 720  
DB 661 ATTGCTTTCTTACGATTTATCAACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA 720  
QY 721 ATTTATCACGTAAAAAGATGAGGGGCCAACAGGATTTTACGCAACACCTTTTAAGGCCCAAT 780  
DB 721 ATTTATCACGTAAAAAGATGAGGGGCCAACAGGATTTTACGCAACACCTTTTAAGGCCCAAT 780  
QY 781 CTAGTGGAAACTAGTTGTCTCCAAACATACCGATGGATCCAAATTTATGCCCGCTGATCCA 840  
DB 781 CTAGTGGAAACTAGTTGTCTCCAAACATACCGATGGATCCAAATTTATGCCCGCTGATCCA 840  
QY 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCTAACCAAGAGAA 900  
DB 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCTAACCAAGAGAA 900  
QY 901 GARAGGAGACTCACTTTATGTCACATGAGCTAGCCACATGGRAGAGATGCCAGATCA 960  
DB 901 GARAGGAGACTCACTTTATGTCACATGAGCTAGCCACATGGRAGAGATGCCAGATCA 960  
QY 961 GCTCGAGTCTGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020  
DB 961 GCTCGAGTCTGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020  
QY 1021 TGTCCAAAGCTGAAACCACTCCAGTGTGTGTAATAATTTCTAGAAATTTCTATATGAAGAA 1080  
DB 1021 TGTCCAAAGCTGAAACCACTCCAGTGTGTGTAATAATTTCTAGAAATTTCTATATGAAGAA 1080  
QY 1081 AATGATGAATCTGAGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAAACCAAGGAG 1140  
DB 1081 AATGATGAATCTGAGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAAACCAAGGAG 1140  
QY 1141 GCAGATTCATTCTCCGTTATCTTAAACAGGTACAGCAAAATCCTTAAAGTCACCTTTTAAAGATG 1200  
DB 1141 GCAGATTCATTCTCCGTTATCTTAAACAGGTACAGCAAAATCCTTAAAGTCACCTTTTAAAGATG 1200  
QY 1201 TCAGAAATAATGATGATGCTGAAAGGAAGATCATAGATGTCTAGATAGAGAACTAAT 1260  
DB 1201 TCAGAAATAATGATGATGCTGAAAGGAAGATCATAGATGTCTAGATAGAGAACTAAT 1260  
QY 1261 CAACCTTTTGAGATTTCTATTGAAAGGAGTTGAATATATTGCGAGAGCTGGATGACTCCT 1320  
DB 1261 CAACCTTTTGAGATTTCTATTGAAAGGAGTTGAATATATTGCGAGAGCTGGATGACTCCT 1320  
QY 1321 GAGGAAATATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTG 1380  
DB 1321 GAGGAAATATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTG 1380  
QY 1381 TTGATCTCACTGAAATTTATTTTCCAGTGAAGATGATGTTTGGAAAGCGAGACTC 1440  
DB 1381 TTGATCTCACTGAAATTTATTTTCCAGTGAAGATGATGTTTGGAAAGCGAGACTC 1440  
QY 1441 ATTGAGTCACTGCGTATTTCTGAGCGCATTAATTTCTATGAAAGAACACAGACTC 1500  
DB 1441 ATTGAGTCACTGCGTATTTCTGAGCGCATTAATTTCTATGAAAGAACACAGACTC 1500  
QY 1501 TGGATAAATCCATGACATCTTTTCTGTTTCCCAAAGTCCGAAAGAGAAATTTGAG 1560  
DB 1501 TGGATAAATCCATGACATCTTTTCTGTTTCCCAAAGTCCGAAAGAGAAATTTGAG 1560  
QY 1561 TTTATTTTGGCTCTGAAATGCAAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620  
DB 1561 TTTATTTTGGCTCTGAAATGCAAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620

Qy	1621	TTAAGGAAAGCAATATAAAGCATCCAGTGGTGGCTGCTCCTCAAGTGATTTCAAG	1680
Db	1621	TTAAGGAAAGCAATATAAAGCATCCAGTGGTGGCTGCTCCTCAAGTGATTTCAAG	1680
Qy	1681	TGTCCTATCAAGAGGAGATAGCAATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAGAGGAGATAGCAATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCAAGTGGTGAATAGTCAAGAGGCTGGTATATTTGAAGGCCACCAAGAC	1800
Db	1741	GGATCTAATATCAAGTGGTGAATAGTCAAGAGGCTGGTATATTTGAAGGCCACCAAGAC	1800
Qy	1801	TCCCCCTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCCTGGAGAGGTGACAAGG	1860
Db	1801	TCCCCCTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCCTGGAGAGGTGACAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGCTGCATCAGTCAGCACTGTGTATTTCTTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACATTTCTTGTGCTGCATCAGTCAGCACTGTGTATTTCTTTTATA	1920
Qy	1921	AGTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCCTTTACAAGCTATCAAGTCCT	1980
Db	1921	AGTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCCTTTACAAGCTATCAAGTCCT	1980
Qy	1981	GAAGATGACCCCACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGT	2040
Db	1981	GAAGATGACCCCACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTATACA	2100
Db	2041	CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTATACA	2100
Qy	2101	TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAGAAATATCCTACTGTG	2160
Db	2101	TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAGAAATATCCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTGTGTC	2179
Db	2161	CTGTTTCATATATGTGTGTC	2179


Search completed: May 4, 2006, 05:19:47  
Job time : 1778 secs

**THIS PAGE BLANK (USPTO)**


GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 4, 2006, 05:20:06 ; Search time 11795 Seconds  
(without alignments)  
123776.066 Million cell updates/sec

Title: US-10-825-632-2  
Perfect score: 3120  
Sequence: 1 aagtgttaagcttcgagg.....agaattactaaaaaaa 3120

Scoring table:    
Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size:    
Total number of hits satisfying chosen parameters: 82156240

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:\*

- 1: gb\_est1.\*
- 2: gb\_est2.\*
- 3: gb\_est3.\*
- 4: gb\_est4.\*
- 5: gb\_est5.\*
- 6: gb\_est6.\*
- 7: gb\_est7.\*
- 8: gb\_est8.\*
- 9: gb\_est9.\*
- 10: gb\_gss2.\*
- 11: gb\_gss3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1664	53.3	2292	4	CR609512 full-length
2	966	31.0	2649	10	AY411615 Homo sapi
3	782	25.1	1292	4	AF175225 Homo sapi
4	769	24.6	1042	3	BM557438 AGENCOURT
5	755	24.2	884	7	CR985650 CR985650
6	726	23.3	910	2	BG479035 602526111
7	725	23.2	985	3	BQ068650 AGENCOURT
8	724	23.2	968	5	BQ671635 AGENCOURT
9	723	23.2	910	5	BQ675260 AGENCOURT
10	722	23.1	1265	5	AF176779 Homo sapi
11	708	22.7	735	1	AL043338 DKFZp4340
12	700	22.4	753	1	AL040398 DKFZp434A
13	690	22.1	835	7	CR998849 CR998849
14	687	22.0	742	7	CN427229 170006002
15	686	22.0	914	5	BX390898 BX390898
16	685	22.0	724	7	CK000768 AGENCOURT
17	682	21.9	820	7	CR994557 CR994557
18	676	21.7	864	1	AL542617 AL542617
19	674	21.6	2649	10	AY411616 Pan trogl
20	672	21.5	1002	2	BE733691 601569288
21	671	21.5	671	8	CX787143 HESC3.81
22	663	21.2	804	5	BU687687 UI-CF-EC1

3/65

ALIGNMENTS

RESULT 1  
CR609512  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
ORIGIN

2292 bp mRNA linear HTC 21-JUL-2004  
full-length cDNA clone CS0DL005YD02 of B cells (Ramos cell line)  
Cot 25-normalized of Homo sapiens (human).

CR609512  
CR609512.1 GI:50490319  
HTC; CDSLT\_CDNA.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 2292)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
2 (bases 1 to 2292)  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
Location/Qualifiers  
1. .2292  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="CS0DL005YD02"  
/tissue\_type="B cells (Ramos cell line) Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"

Query Match 53.3%; Score 1664; DB 4; Length 2292;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1714; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

23	660	21.2	993	2	BG259714	BG259714 602380072
c	652	20.9	910	5	BX372276	BX372276 BX372276
24	647	20.7	789	2	BG709118	BG709118 602675382
25	644	20.6	669	2	BG390103	BG390103 602415944
26	641	20.5	881	7	CR990494	CR990494 CR990494
27	639	20.5	708	5	BG222228	BG222228 UI-H-FH1-
c	637	20.4	722	5	BU631054	BU631054 UI-H-FH1-
c	627	20.1	853	2	BI223892	BI223892 602941035
30	626	20.1	815	6	CB990233	CB990233 AGENCOURT
31	626	20.1	826	3	BI553405	BI553405 603193351
32	623	20.0	695	2	BG715866	BG715866 602675282
33	623	20.0	634	1	AI819365	AI819365 wg61c02.x
c	614	19.7	634	1	AI819365	AI819365 wg61c02.x
35	607	19.5	957	5	BQ675006	BQ675006 AGENCOURT
36	598	19.2	612	7	CV024163	CV024163 1359 Full
37	596	19.1	596	6	CB153590	CB153590 K-EST0211
38	581	18.6	632	5	BX645431	BX645431 DKFZp781O
39	580	18.6	1041	2	BI084090	BI084090 602869453
40	576	18.5	627	2	BG752328	BG752328 602730802
c	562	18.0	760	1	AI917735	AI917735 t11d07.x
41	561	18.0	678	7	CN427235	CN427235 170005332
c	558	17.9	673	1	AW303607	AW303607 xv21b12.x
43	555	17.8	788	2	BG163397	BG163397 602338360
44	548	17.6	553	7	CN427230	CN427230 170005332

1/03

472 GCCATGCTGCTGAGAACACAGAAATACACTGTTTATTTCTGAAATCCCAAACTATC 531  
11 GCCATGCTGCTGAGAACACAGAAATACACTGTTTATTTCTGAAATCCCAAACTATC 70  
532 AATAGACGACGAGCTTAAATGCTCTCTCGAAGCCCTCTTTGGATCTTTTTCAGGCAACA 591  
Db AATAGACGACGAGCTTAAATGCTCTCTCGAAGCCCTCTTTGGATCTTTTTCAGGCAACA 130  
592 CTGACATATGGAATGTTATCTCGAAGAAAGAACTATTAAGAGAAAGAAACGCAATTGGA 651  
Db CTGACATATGGAATGTTATCTCGAAGAAAGAACTATTAAGAGAAAGAAACGCAATTGGA 190  
652 ACAGTCGGAATGCTCTTACGATATCAACGAAGAGTGGAACTTTCTGTTTCAAGCC 711  
Db ACAGTCGGAATGCTCTTACGATATCAACGAAGAGTGGAACTTTCTGTTTCAAGCC 250  
712 GGTAGTGGAAATTTATCACTGTAAGAGTGGAGGCCACAAGGATTTTACGCAACAACTTTTA 771  
Db GGTAGTGGAAATTTATCACTGTAAGAGTGGAGGCCACAAGGATTTTACGCAACAACTTTTA 310  
772 AGGCCCAATCTAGTGGAAATCTAGTTGTCCTCAACATACGATGGATGCCAAATATGCCCC 831  
Db AGGCCCAATCTAGTGGAAATCTAGTTGTCCTCAACATACGATGGATGCCAAATATGCCCC 370  
832 GCTGATCCAGACTGGATTTGCTTTTATACATAGCAAGATATTTGGATATCTAAACATGTA 891  
Db GCTGATCCAGACTGGATTTGCTTTTATACATAGCAAGATATTTGGATATCTAAACATGTA 430  
892 ACCAGAGAAAGAAAGGAGACTCACTATATGTCACATAGCTAGCCCAACATCGAAGAGAT 951  
Db ACCAGAGAAAGAAAGGAGACTCACTATATGTCACATAGCTAGCCCAACATCGAAGAGAT 490  
952 GCCAGATCAGCTGGAGTGGCTACCTTTGTCCTCAAGAGAAATTTGATAGATTTCTGGC 1011  
Db GCCAGATCAGCTGGAGTGGCTACCTTTGTCCTCAAGAGAAATTTGATAGATTTCTGGC 550  
1012 TATTTGGTGGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATCTTA 1071  
Db TATTTGGTGGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTAGAAATCTTA 610  
1072 TATGAAGAAATGATGAATCTGAGTGGAAATTTATTCATGTTACATCCCTTATGTTGAA 1131  
Db TATGAAGAAATGATGAATCTGAGTGGAAATTTATTCATGTTACATCCCTTATGTTGAA 670  
1132 ACAAGGAGGAGATTCATTCCTGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACT 1191  
Db ACAAGGAGGAGATTCATTCCTGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACT 730  
1192 TTTAAGATGTCAGAAATTAATGATGCTGTAAGGAAGATCATAGATGTCATAGATAAG 1251  
Db TTTAAGATGTCAGAAATTAATGATGCTGTAAGGAAGATCATAGATGTCATAGATAAG 790  
1252 GAACATAATCAACCTTTGAGATTTCTATTTGAAGGAGTTGAATATTTGCCAGAGCTGGA 1311  
Db GAACATAATCAACCTTTGAGATTTCTATTTGAAGGAGTTGAATATTTGCCAGAGCTGGA 850  
1312 TGGACTCCTCAGGGAATAATGCTGGTCCATCTCTAGATAGCTCCAGACTCGCCTA 1371  
Db TGGACTCCTCAGGGAATAATGCTGGTCCATCTCTAGATAGCTCCAGACTCGCCTA 910  
1372 CAGATAGTGTGATCTCAGCTGAATTTATCCAGTGAAGATGATGTTATGGAAG 1431  
Db CAGATAGTGTGATCTCAGCTGAATTTATCCAGTGAAGATGATGTTATGGAAG 970  
1432 CAGAGACTATTCAGTCACTGCTGATTTCTGTCGCACTAATATCTATGGAAGAAACA 1491  
Db CAGAGACTATTCAGTCACTGCTGATTTCTGTCGCACTAATATCTATGGAAGAAACA 1030  
1492 ACAGACATCTGGATAAATATCAATGATCTTTTCAATGTTTTTCCCAAGTCAAGAGAG 1551  
Db ACAGACATCTGGATAAATATCAATGATCTTTTCAATGTTTTTCCCAAGTCAAGAGAG 1090  
1552 GAAATGAGTTTATTTTCCCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAATTT 1611

Db 1091 GAAATGAGTTTATTTTCCCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAATTT 1150  
Qy 1612 ACATCTATTTTAAAGGAAGCAAAATATAAAGCATCAAGTCAAGTGGGCTGCTGCTCCAAAGT 1671  
Db 1151 ACATCTATTTTAAAGGAAGCAAAATATAAAGCATCAAGTCAAGTGGGCTGCTGCTCCAAAGT 1210  
Qy 1672 GATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTT 1731  
Db 1211 GATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTT 1270  
Qy 1732 GCGCGGATCGATCTAATATCAAGTGTGAAGTCAAGGCTCAGAAAGCTGCTATATTTTGAAGGC 1791  
Db 1271 GCGCGGATCGATCTAATATCAAGTGTGAAGTCAAGGCTCAGAAAGCTGCTATATTTTGAAGGC 1330  
Qy 1792 ACCAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAAATCCCTGGAGAG 1851  
Db 1331 ACCAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAAATCCCTGGAGAG 1390  
Qy 1852 GTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCACTGTGAC 1911  
Db 1391 GTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCACTGTGAC 1450  
Qy 1912 TTCTTTATAGTAAAGTATAGTAACAGAGAGATCCACACTGTGTGCTCCCTTTCACAGCTA 1971  
Db 1451 TTCTTTATAGTAAAGTATAGTAACAGAGAGATCCACACTGTGTGCTCCCTTTCACAGCTA 1510  
Qy 1972 TCAAGTCTCGAAGTCAACCCAACTTGCACAAACAAAGAAATTTGGGCCACCACTTTGGAT 2031  
Db 1511 TCAAGTCTCGAAGTCAACCCAACTTGCACAAACAAAGAAATTTGGGCCACCACTTTGGAT 1570  
Qy 2032 TCAGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACT 2091  
Db 1571 TCAGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACT 1630  
Qy 2092 GGATTTACATGTTATGGGAGTCTACAAAGCTCAATGATCTACAGCTCGGAAAGAAATAT 2151  
Db 1631 GGATTTACATGTTATGGGAGTCTACAAAGCTCAATGATCTACAGCTCGGAAAGAAATAT 1690  
Qy 2152 CCTACTGTGCTGTTTCATATATATGTTGCTCTCAGGT 2186  
Db 1691 CCTACTGTGCTGTTTCATATATGTTGCTCTCAGGT 1725

RESULT 2  
AY411615  
LOCUS  
DEFINITION Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,  
Genomic survey sequence.  
ACCESSION  
AY411615  
VERSION  
AY411615.1 GI:39767583  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo (base 1 to 2649)  
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
Science 302 (5652), 1960-1963 (2003)  
JOURNAL  
PUBMED  
14671302  
REFERENCE  
2. (bases 1 to 2649)  
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.

FEATURES  
source  
1..2649  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>2649  
/gene="DPP8"  
/locus\_tag="HCM4283"

ORIGIN  
Query Match 31.0%; Score 966; DB 10; Length 2649;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1016; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 214 ATGGCAGCAGCAATGAAACAGACAGAGCTGGGTGTGAGATATTTGAAACTGCGACTGT 273  
DB 1 ATGGCAGCAGCAATGAAACAGACAGAGCTGGGTGTGAGATATTTGAAACTGCGACTGT 60

QY 274 GAGGAGAAATTAATCAAGATCGGCTTAAATTTGAGAGCCTTTTATGTTGAGCGGTAT 333  
DB 61 GAGGAGAAATTAATCAAGATCGGCTTAAATTTGAGAGCCTTTTATGTTGAGCGGTAT 120

QY 334 TCCTGAGTCAGCTTAAAGCTGCTGCCGATACCAAGAAATATCATGCTACATGATG 393  
DB 121 TCCTGAGTCAGCTTAAAGCTGCTGCCGATACCAAGAAATATCATGCTACATGATG 180

QY 394 GCTAAGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATCA 453  
DB 181 GCTAAGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATCA 240

QY 454 GACAGAACTATTACCTGCCATGCTGGTGAGAACAGAGAAATACACTGTTTATTCT 513  
DB 241 GACAGAACTATTACCTGCCATGCTGGTGAGAACAGAGAAATACACTGTTTATTCT 300

QY 514 GAAATCCCAAACTATCAATAGCAGCAGCTTAAATGCTCTCTCGAAGACCTCTTTTG 573  
DB 301 GAAATCCCAAACTATCAATAGCAGCAGCTTAAATGCTCTCTCGAAGACCTCTTTTG 360

QY 574 GATCTTTTTCAGGCAACACTGGAATGATGATTTCTCGAAGAGAACTATTAAGA 633  
DB 361 GATCTTTTTCAGGCAACACTGGAATGATGATTTCTCGAAGAGAACTATTAAGA 420

QY 634 GAAAGAAAACCGCATTTGGAACAGTCGGAATGCTTTACGATTTACCAAGGAAGTGA 693  
DB 421 GAAAGAAAACCGCATTTGGAACAGTCGGAATGCTTTACGATTTACCAAGGAAGTGA 480

QY 694 AATTTCTGTTCAAGCCGCTAGTGAATTTATCAGTAAAGATGAGGAGCCCAAGGA 753  
DB 481 AATTTCTGTTCAAGCCGCTAGTGAATTTATCAGTAAAGATGAGGAGCCCAAGGA 540

QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACCTAGTTGTCCCAACATACGATG 813  
DB 541 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACCTAGTTGTCCCAACATACGATG 600

QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATTT 873  
DB 601 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATTT 660

QY 874 TGGATATTAACATCGTAACCAAGAGAAAGGAGACTCACTTATGTGCAAAATGAGCTA 933  
DB 661 TGGATATTAACATCGTAACCAAGAGAAAGGAGACTCACTTATGTGCAAAATGAGCTA 720

QY 934 GCCAACATGGAAGAAATGCGAGATCAGCTGGAGTCCGTACCTTTGTTCTCCAAAGAGAA 993  
DB 721 GCCAACATGGAAGAAATGCGAGATCAGCTGGAGTCCGTACCTTTGTTCTCCAAAGAGAA 780

QY 994 TTTGATAGATATTTGCTGCTATTGGTGGTGTCCAAAAGCTGAAAACAACTCCCAAGTGGT 1053  
DB 781 TTTGATAGATATTTGCTGCTATTGGTGGTGTCCAAAAGCTGAAAACAACTCCCAAGTGGT 840

QY 1054 AAAATCTTAGAATTTCTATATGAAGAAATGATGATCTGAGGTGAAATATTTCATGTT 1113  
DB 841 AAAATCTTAGAATTTCTATATGAAGAAATGATGATCTGAGGTGAAATATTTCATGTT 900

QY 1114 ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCATTCGTTATCTCTAAACAGGTACA 1173  
DB 901 ACATCCCTATGTTGGAACCAAGGAGGCGAGATTCATTCGTTATCTCTAAACAGGTACA 960

QY 1174 GCAAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATGCTGAAGAAGG 1230  
DB 961 GCAAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATGCTGAAGAAGG 1017

RESULT 3  
AF175225 1292 bp mRNA linear HTC 01-AUG-2003  
LOCUS Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.  
DEFINITION AF175225  
ACCESSION AF175225  
VERSION AF175225.1 GI:33338055  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1292)  
AUTHORS Zhao, B., Xu, H.S., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B.,  
Wang, X.Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J.,  
Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zheng, W.Y., Teng, C.Y.,  
Liu, Q., Yu, D.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hui, R.T.  
Direct Submission  
Submitted (04-AUG-1999) Molecular Medicine Center for  
Cardiovascular Disease, Cardiovascular Institute, CNMS & PUMC, 167,  
Bei Li Shi Lu, Beijing 100037, P.R. China  
FEATURES  
source  
1..1292  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="aorta"  
142..1023  
/codon\_start=1  
/product="MSTP135"  
/protein\_id="AAQ13650.1"  
/db\_xref="GI:33338056"  
/translation="MTQAKORNFPPFWIQVLFLLILKQFSLKVLKLLDLMCMGCS  
TSPHSGTGLERNILLXLFYGPQKVLVNNRFGKVFRLNLPFLGVVXGIDNRGS  
CPKGLFEGAFKFMQGIIDQVEGLQYLSRVDYIDLDRVGIHWSYGYLSLML  
MQRSDIFRVAIAGAVTLLWIFDTGTERYMGHPDQNEQGYILGSVAMQAKPSPE  
RLLLHGFIDENVHFAHTSILLSFLVRAGKPYDILQIYPQERKHSIRVPESGHEVHLL  
HYLQENLGSRIAALKVI"  
ORIGIN  
Query Match 25.1%; Score 784; DB 4; Length 1292;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2325 TAAATGGGCAAAATAGAAATTCAGCATCAGGTGGAAGACTCCAAATATCTAGCTTCTCG 2384  
DB 486 TAAATGGGCAAAATAGAAATTCAGCATCAGGTGGAAGACTCCAAATATCTAGCTTCTCG 545

QY 2385 ATATGATTTTCATTTAGATCGTGTGCGCATCCACGGCTGCTTATGAGGATACCT 2444  
DB 546 ATATGATTTTCATTTAGATCGTGTGCGCATCCACGGCTGCTTATGAGGATACCT 605

QY 2445 CTCCTGATGGCATTATGACAGAGTCAGATATCTTCAGGTTGCTATTGCTGGGGCCCC 2504  
DB 606 CTCCTGATGGCATTATGACAGAGTCAGATATCTTCAGGTTGCTATTGCTGGGGCCCC 665

QY 2505 AGTCACCTCTGGATCTTCTATGATACAGGATACACGGACCTTATATGGGTACCCCTGA 2564  
DB 666 AGTCACCTCTGGATCTTCTATGATACAGGATACACGGACCTTATATGGGTACCCCTGA 725

```
QY 2565 CCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTC 2624
Db 726 CCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTC 785
QY 2625 TGAACCAAAATCGTTTACTGCTCTACATGTTTCTCGGATGAGATGCAATTTTGACA 2684
Db 786 TGAACCAAAATCGTTTACTGCTCTACATGTTTCTCGGATGAGATGCAATTTTGACA 845
QY 2685 TACCAGTATATTACTGATGTTTATGATGAGGCTGGAAGCCATATGATTTACAGATCTA 2744
Db 846 TACCAGTATATTACTGATGTTTATGATGAGGCTGGAAGCCATATGATTTACAGATCTA 905
QY 2745 TCCTCAGAGAGACACAGCATTAAGAGTTCTGAAATCGGAGAACATTAAGATGCAATCT 2804
Db 906 TCCTCAGAGAGACACAGCATTAAGAGTTCTGAAATCGGAGAACATTAAGATGCAATCT 965
QY 2805 TTTGCATCTCTCAAGAAAACCTTGGATCAGCTATTGCTCTTAAAGTGATATAATT 2864
Db 966 TTTGCATCTCTCAAGAAAACCTTGGATCAGCTATTGCTCTTAAAGTGATATAATT 1025
QY 2865 TTGACCTGTGTGAAGCTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTTTAAAT 2924
Db 1026 TTGACCTGTGTGAAGCTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTTTAAAT 1085
QY 2925 CAACAGAAAACACAGAAATGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCCT 2984
Db 1086 CAACAGAAAACACAGAAATGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCCT 1145
QY 2985 GAAATAAATGTGGTCCCATGACAGGGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAA 3044
Db 1146 GAAATAAATGTGGTCCCATGACAGGGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAA 1205
QY 3045 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCAATAGAA 3104
Db 1206 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCAATAGAA 1265
QY 3105 TT 3106
Db 1266 TT 1267
```

```
RESULT 4
BM557438
LOCUS BM557438 1042 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_578992 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466759
S'; mRNA Sequence.
ACCESSION BM557438
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1042)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: DCTD/DPB
cDNA library preparation: Rubin Laboratory
cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.lln.gov
Plate: L1042
High quality sequence stop: 697.
Location/Qualifiers
1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
```

FEATURES  
source

```
/db_xref="taxon:9606"  
/clone="IMAGE:5466759"  
/tissue_type="amelanotic melanoma, cell line"  
/lab_host="DH10B (phage-resistant)"  
/clone_lib="NIH MGC 41"  
/notes="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH_MGC Library."
```

## ORIGIN

```
Query Match 24.6%; Score 769; DB 3; Length 1042;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 819; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1723 GAAGTCTTGGCCGGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATAT 1782  
Db 1 GAAGTCTTGGCCGGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATAT 60  
QY 1783 TTTGAGGGACCAAAAGCTCCCTTTAGAGCATACCTGTAGTACGATTCAGTTCAGTAAAT 1842  
Db 61 TTTGAGGGACCAAAAGCTCCCTTTAGAGCATACCTGTAGTACGATTCAGTTCAGTAAAT 120  
QY 1843 CCTGGAGAGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAG 1902  
Db 121 CCTGGAGAGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAG 180  
QY 1903 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTT 1962  
Db 181 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTT 240  
QY 1963 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTTGCAAAAACAAAGGAATTTGGGGCACC 2022  
Db 241 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTTGCAAAAACAAAGGAATTTGGGGCACC 300  
QY 2023 ATTTTGGATTACAGAGGTCCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTGAA 2082  
Db 301 ATTTTGGATTACAGAGGTCCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTGAA 360  
QY 2083 AGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCTCATATCTACAGCCTGGA 2142  
Db 361 AGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCTCATATCTACAGCCTGGA 420  
QY 2143 AAGAAATATCCTACTGTCGCTGTTCATATATGTTGCTCCTCAGGTGCAGTTGGTGAATAT 2202  
Db 421 AAGAAATATCCTACTGTCGCTGTTCATATATGTTGCTCCTCAGGTGCAGTTGGTGAATAT 480  
QY 2203 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGNATACCTAGCTCTCTAGTTATGGTTT 2262  
Db 481 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGNATACCTAGCTCTCTAGTTATGGTTT 540  
QY 2263 GTAGTGATAGACACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGCGCCTTTAAA 2322  
Db 541 GTAGTGATAGACACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGCGCCTTTAAA 600  
QY 2323 TATAAAATGGGTCAAAATAGAAAATTCAGATCAGGTGGAAGGACTCCAAATATCTAGTTCT 2382  
Db 601 TATAAAATGGGTCAAAATAGAAAATTCAGATCAGGTGGAAGGACTCCAAATATCTAGTTCT 660  
QY 2383 CGATATGATTTTCAATGACTTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGATAC 2442  
Db 661 CGATATGATTTTCAATGACTTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGATAC 720  
QY 2443 CTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTACAGGGTTCCTATTGCTGGGGC 2502  
Db 721 CTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTACAGGGTTCCTATTGCTGGGGC 780  
QY 2503 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGG 2542
```





```
Plate: LLCMI430 row: e column: 03
High quality sequence stop: 837.
Location/Qualifiers
1. .910
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:464934"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/notes="Organ: Placenta; Vector: pOTB7; Site: 1: XhoI; Site_2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

Query Match 23.3%; Score 726; DB 2; Length 910;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2193 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGG 2252  
Db 1 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGG 60

QY 2253 TTATGTGTTGTAGTGATAGACAAACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAGG 2312  
Db 61 TTATGTGTTGTAGTGATAGACAAACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAGG 120

QY 2313 CGCCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCAGGTGAAGGACTCCAAAT 2372  
Db 121 CGCCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCAGGTGAAGGACTCCAAAT 180

QY 2373 TCTAGCTTCGAGATATGATTCATTTGACTTAGATCGTGTGGCATCAACGGCTGTCTTA 2432  
Db 181 TCTAGCTTCGAGATATGATTCATTTGACTTAGATCGTGTGGCATCAACGGCTGTCTTA 240

QY 2433 TGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTAT 2492  
Db 241 TGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTAT 300

QY 2493 TGTGGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATAT 2552  
Db 301 TGTGGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATAT 360

QY 2553 GGGTACCCCTGACCAAGATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGA 2612  
Db 361 GGGTACCCCTGACCAAGATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGA 420

QY 2613 AAAAGTCCCTCTGAAACCAATCGTTTACTGCTTTACATGTTTCTTGGATGAGAAATGT 2672  
Db 421 AAAAGTCCCTCTGAAACCAATCGTTTACTGCTTTACATGTTTCTTGGATGAGAAATGT 480

QY 2673 CCATTTTGACATACACAGTATATTAATCTAGTTTTTTTATGAGGGCTGGAAGCCATATCA 2732  
Db 481 CCATTTTGACATACACAGTATATTAATCTAGTTTTTTTATGAGGGCTGGAAGCCATATCA 540

QY 2733 TTTACAGATCTATCTTCAGGAGACACAGCATATAGAGTTCCCTGAATCGGAGAACATTA 2792  
Db 541 TTTACAGATCTATCTTCAGGAGACACAGCATATAGAGTTCCCTGAATCGGAGAACATTA 600

QY 2793 TGAACATGATCTTTTGCACTACCTTCAGRAAACCCTTGGATTCAGTATTCGCTCTTAA 2852  
Db 601 TGAACATGATCTTTTGCACTACCTTCAGRAAACCCTTGGATTCAGTATTCGCTCTTAA 660

QY 2853 AGTGATATAATTTTGAACCTGTGAGAACTCTCTGGTATACATCGGCTATTTTAAACCAATG 2912  
Db 661 AGTGATATAATTTTGAACCTGTGAGAACTCTCTGGTATACATCGGCTATTTTAAACCAATG 720

QY 2913 AGGAGG 2918  
Db 721 AGGAGG 726

RESULT 7  
BQ068650  
LOCUS BQ068650.1 GI:19897696  
DEFINITION 5', mRNA sequence.  
ACCESSION BQ068650  
VERSION BQ068650.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 985)  
NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2035 row: 0 column: 08  
High quality sequence stop: 677.  
Location/Qualifiers  
1. .985  
/organism="Homo sapiens"  
/mol\_type="mrna"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5802079"  
/tissue\_type="neuroblastoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_47"  
/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 23.2%; Score 725; DB 3; Length 985;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TTCGCCGCTGGGTTGTACCGGGCCGCCCGGAGGAGCCACTGCAACGAGGCCGA 122  
Db 1 TTCGCCGCTGGGTTGTACCGGGCCGCCCGGAGGAGCCACTGCAACGAGGCCGA 60

QY 123 GTGAGGCGCGGCGAGCATGACCGCGCGCCGCTCCATAGCGCAGTCGCGGACGGTC 182  
Db 61 GTGAGGCGCGGCGAGCATGACCGCGCGCCGCTCCATAGCGCAGTCGCGGACGGTC 120

QY 183 CGGCGCGCGCGCGGCGGAGGAAATGCAATCGCAGCAGCAATGGAACAGAACGCT 242  
Db 121 CGGCGCGCGCGCGGCGGAGGAAATGCAATCGCAGCAGCAATGGAACAGAACGCT 180

QY 243 GGGTGTGTAGATATTTGAAACTTGGGACTGTGAGGAGAAATTTGAATTCACAGGATCGGCC 302  
Db 181 GGGTGTGTAGATATTTGAAACTTGGGACTGTGAGGAGAAATTTGAATTCACAGGATCGGCC 240



ACCESSION	BO675260	BO675260.1	GI:21786094			
VERSION	EST..					
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.					
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .					
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLCM2458 row: 1 column: 07 High quality sequence stop: 618.					
FEATURES	Location/Qualifiers					
source	1..910					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/clones="IMAGE:6275478"					
	/tissue_type="epidermoid carcinoma, cell line"					
	/lab_host="DHIOB (phage-resistant)"					
	/clone_lib="NIH_MGC_102"					
	/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."					
ORIGIN						
Query Match	23.2%;	Score 723;	DB 5;	Length 910;		
Best Local Similarity	99.8%;	Pred. No. 0;				
Matches 823;	Conservative	0;	Mismatches	2;		
	Indels	0;	Gaps	0;		
QY	1635	ATATAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGATTTCAAGTGTCCTATCAAGA	1694			
Db	1	ATATAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGATTTCAAGTGTCCTATCAAGA	60			
QY	1695	GGAGATAGCAATTACCAATGCAATGGGAAGTTCTTGGCCGGCATGATCTAATATCCA	1754			
Db	61	GGAGATAGCAATTACCAATGCAATGGGAAGTTCTTGGCCGGCATGATCTAATATCCA	120			
QY	1755	AGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCCCAAGACTCCCTTTAGAGCA	1814			
Db	121	AGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCCCAAGACTCCCTTTAGAGCA	180			
QY	1815	TCACCTGTACGTAGTCAGTACGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGG	1874			
Db	181	TCACCTGTACGTAGTCAGTACGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGG	240			
QY	1875	CTACTCATTCTTGTGTCATAGTCAGTCAGTCATCTTTTAAAGTAAGTATAGTAA	1934			
Db	241	CTACTCATTCTTGTGTCATAGTCAGTCAGTCATCTTTTAAAGTAAGTATAGTAA	300			
QY	1935	CCAGAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGCCCAAC	1994			
Db	301	CCAGAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGCCCAAC	360			
QY	1995	TTGCAAAACAAAGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCGACTA	2054			
Db	361	TTGCAAAACAAAGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCGACTA	420			
QY	2055	TACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGATGCT	2114			
Db	421	TACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGATGCT	480			
QY	2115	CTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGCTGTTTATATATGG	2174			
Db	481	CTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGCTGTTTATATATGG	540			
QY	2175	TGGTCTCAGGTGAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGA	2234			
Db	541	TGGTCTCAGGTGAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGA	600			
QY	2235	TACCTAGCCTCTCTAGGTTATGTGGTTAGTGTAGTATAGACACAGGGATCTGTACCG	2294			
Db	601	TACCTAGCCTCTCTAGGTTATGTGGTTAGTGTAGTATAGACACAGGGATCTGTACCG	660			
QY	2295	AGGCTTTAAATTTGAAGGGCCCTTTAAATATAAAATGGTCAATAGAAAATGACGATCA	2354			
Db	661	AGGCTTTAAATTTGAAGGGCCCTTTAAATATAAAATGGTCAATAGAAAATGACGATCA	720			
QY	2355	GGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGGG	2414			
Db	721	GGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGGG	780			
QY	2415	CATCCAGCGCTGGTCTATGGAGGATACCTTCCCTGATGGCATT	2459			
Db	781	CATCCAGCGCTGGTCTATGGAGGATACCTTCCCTGATGGCATT	825			
RESULT 10						
AFI76779						
LOCUS	AFI76779	Homo sapiens	MSTP141	mRNA linear HTC 01-AUG-2003		
DEFINITION	Homo sapiens MSTP141 mRNA, complete cds.					
ACCESSION	AFI76779					
VERSION	AFI76779.1	GI:33338069				
KEYWORDS	HTC.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 1265)					
AUTHORS	Hui, R. T., Liu, Y. Q., Wang, X. Y., Qin, B. M. and Sheng, H.					
TITLE	Homo sapiens normal aorta mRNA MST141					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 1265)					
AUTHORS	Hui, R. T., Liu, Y. Q., Wang, X. Y., Qin, B. M. and Sheng, H.					
TITLE	Direct Submission					
JOURNAL	Submitted (10-AUG-1999) Molecular Medicine Center for Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167, Bei Li Shi Lu, Beijing 100037, P. R. China					
FEATURES	Location/Qualifiers					
source	1..1265					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/tissue_type="aorta"					
	268..741					
	/codon_start=1					
	/product="MSTP141"					
	/protein_id="AAQ13657.1"					
	/db_xref="GI:3338070"					
	/translation="MLYKPHDLQPGKPYTVLFYGGFQVOLNNRFGVKYFRLNLT ASLGYYVVVIDNRGSGCHRLKFGKYMGOIEIDQVEGLQYLASRYDFILDVRG IHGWSYGGYLSLMLMQRSDIFRVAIAGAPVTLWIFYDTGYTERYMGHPQNEQG"					
ORIGIN						
Query Match	23.1%;	Score 722;	DB 4;	Length 1265;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 722;	Conservative	0;	Mismatches	0;		
	Indels	0;	Gaps	0;		
QY	1861	CTGACTGACCGTGGCTACTCACATCTTGTGTCATCAGTCAGCACTGTGACTTCTTTATA	1920			

Db

19

CTGACTGACCGTGGCTACTCAGATCTTCTGTCATCAGTCAGCACTCTGACTCTTTATA

78

Qy

1921

AGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCT

1980

Db

79

AGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCT

138

Qy

1981

GAGATGACCCAACTTGGCAAAACAAAGGAATTTGGGCCACATTTGGATTGACAGGT

2040

Db

139

GAGATGACCCAACTTGGCAAAACAAAGGAATTTGGGCCACATTTGGATTGACAGGT

198

Qy

2041

CCTCTTCTGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACA

2100

Db

199

CCTCTTCTGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACA

258

Qy

2101

TTGTATGGGATGCTTACAGGCTCATGATCTACAGCTCGAAGGAAATATCTTACTGTG

2160

Db

259

TTGTATGGGATGCTTACAGGCTCATGATCTACAGCTCGAAGGAAATATCTTACTGTG

318

Qy

2161

CTGTTCATATATGTTGGTCTCCTCAGGTCAGTTGGTGAATATCGGTTTAAAGGAGTCAAG

2220

Db

319

CTGTTCATATATGTTGGTCTCCTCAGGTCAGTTGGTGAATATCGGTTTAAAGGAGTCAAG

378

Qy

2221

TATTTCCGTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTGATAGACACAGG

2280

Db

379

TATTTCCGTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTGATAGACACAGG

438

Qy

2281

GGATCTGTCCCGAGGGCTTAAATTTGAAGGCGCTTAAATATAAATGGGTCAATA

2340

Db

439

GGATCTGTCCCGAGGGCTTAAATTTGAAGGCGCTTAAATATAAATGGGTCAATA

498

Qy

2341

GAAATTCAGCATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGAC

2400

Db

499

GAAATTCAGCATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGAC

558

Qy

2401

TTAGATCGTGGGCAATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATT

2460

Db

559

TTAGATCGTGGGCAATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATT

618

Qy

2461

ATGCAGAGGTGAGATATCTTACGGTTCCTATGCGGGGCCACGTCACCTCTGTGATC

2520

Db

619

ATGCAGAGGTGAGATATCTTACGGTTCCTATGCGGGGCCACGTCACCTCTGTGATC

678

Qy

2521

TTCTATGATACAGATACAGGAACCTTATGTTGGTTCACCTGACCAAGATGAACAGGC

2580

Db

679

TTCTATGATACAGATACAGGAACCTTATGTTGGTTCACCTGACCAAGATGAACAGGC

738

Qy

2581

TA 2582

Db

739

TA 740

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by LMU (Ludwig Maximilians University, Munich/Germany) within the cDNA sequencing consortium of the German Genome Project.

No s1 sequence available.

This clone (DKFZp43400723) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1. .735

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp43400723"

/tissue\_type="testis"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="434 (synonym: htes3)"

/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"

FEATURES		source
ORIGIN		
Query Match 22.7%; Score 708; DB 1; Length 735;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	2274	CAACAGGGGATCCTGTCCAGCGGGCTTAAATTTGAAGCGCCCTTTAAATAATAAATGGG 2333
Db	1	CAACAGGGGATCCTGTCCAGCGGGCTTAAATTTGAAGCGCCCTTTAAATAATAAATGGG 60
Qy	2334	TCAAATAGAAATGACGATCAGGTGAAGGATCCAAATATCTAGCTTCTCGATGATTT 2393
Db	61	TCAAATAGAAATGACGATCAGGTGAAGGATCCAAATATCTAGCTTCTCGATGATTT 120
Qy	2394	CATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGAT 2453
Db	121	CATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGAT 180
Qy	2454	GGCATTATACAGAGGTGAGATATCTTCAAGGGTTCCTATGCTGGGGCCCGAGTCACCTCT 2513
Db	181	GGCATTATACAGAGGTGAGATATCTTCAAGGGTTCCTATGCTGGGGCCCGAGTCACCTCT 240
Qy	2514	GTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTGACCTGACCAAGATGA 2573
Db	241	GTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTGACCTGACCAAGATGA 300
Qy	2574	ACAGGGCTATTAATCTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACAAA 2633
Db	301	ACAGGGCTATTAATCTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACAAA 360
Qy	2634	TCGTTTACTGCTCTTACATGGTTTCTGATGAGATGTCATTTTGCACATACCATGAT 2693
Db	361	TCGTTTACTGCTCTTACATGGTTTCTGATGAGATGTCATTTTGCACATACCATGAT 420
Qy	2694	ATTACTCAGTTTTTACTGAGGGCTGAAAGCCATATGATTTTACAGATCTATCTCCAGGA 2753
Db	421	ATTACTCAGTTTTTACTGAGGGCTGAAAGCCATATGATTTTACAGATCTATCTCCAGGA 480
Qy	2754	GAGACAGCATTAAGAGTTCCTGAATCGGAGAAACATTATGAATGCGATCTTTTGCACATA 2813
Db	481	GAGACAGCATTAAGAGTTCCTGAATCGGAGAAACATTATGAATGCGATCTTTTGCACATA 540
Qy	2814	CCTTCAAGAAAACTTGGATCAGCTATTTCTGCTCTTAAAGGTGATATAATTTTGACCTGT 2873
Db	541	CCTTCAAGAAAACTTGGATCAGCTATTTCTGCTCTTAAAGGTGATATAATTTTGACCTGT 600
Qy	2874	GTAGAACTCTCTGTGATACACTGGCTATTTAAACAAATCAGGAGGTTTAATCAACAGAAA 2933
Db	601	GTAGAACTCTCTGTGATACACTGGCTATTTAAACAAATCAGGAGGTTTAATCAACAGAAA 660
Qy	2934	ACACAGAAATTGATCATCATTTTGTGATCTGCTGCAATGTAACATCTACT 2981
Db	661	ACACAGAAATTGATCATCATTTTGTGATCTGCTGCAATGTAACATCTACT 708

RESULT 11

AL043338

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

735 bp

mRNA

linear

EST 04-SEP-2003

DKFZp43400723 r1 434 (synonym: htes3) Homo sapiens cDNA clone

DKFZp43400723-5', mRNA sequence.

AL043338

AL043338.1

GI:5422728

EST

EST

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 735)

Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

EST (Blum, et al.)

Unpublished (1999)

Contact: MIPS

MIPS

Ingolisaedter Landstr. 1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

```

RESULT 12
AL040398
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 753)
Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
EST (Koehrer, et al.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp34A0714) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..753
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp34A0714"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="pH10B"
/clone_lib="434 (synonym: htes3)"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"

FEATURES
source
Query Match 22.4%; Score 700; DB 1; Length 753;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 750; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1609 ATTACATCTATTTTAAAGAAAGCAAAATATAAAGATCCAGTGTGGCTGCCTGCTCCA 1668
Db 1 ATTACATCTATTTTAAAGAAAGCAAAATATAAAGATCCAGTGTGGCTGCCTGCTCCA 60

Qy 1669 AGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTACCAAGTGTGAATGGGAAGTT 1728
Db 61 AGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTACCAAGTGTGAATGGGAAGTT 120

Qy 1729 CTTGGCCGGCATGGATCTAATATCAAGTGTGAAGTCAAGAGGCTGGGTATATTTTGA 1788
Db 121 CTTGGCCGGCATGGATCTAATATCAAGTGTGAAGTCAAGAGGCTGGGTATATTTGA 180

Qy 1789 GGCACAAAGACTCCCTTTAGAGCATCACTGTACGTACGTACGTACGTACGTACGTAC 1848
Db 181 GGCACAAAGACTCCCTTTAGAGCATCACTGTACGTACGTACGTACGTACGTACGTAC 240

Qy 1849 GAGGTGACAAAGCTGACTGACCGTGGCTACTACATCTTGTGTCGATCAGTCAGCTGT 1908
Db 241 GAGGTGACAAAGCTGACTGACCGTGGCTACTACATCTTGTGTCGATCAGTCAGCTGT 300

Qy 1909 GACTCTTTTAAAGTATAGTATACCAAGAAATCCACATGTGTGTGTCCTTTACAG 1969
Db 301 GACTCTTTTAAAGTATAGTATACCAAGAAATCCACATGTGTGTGTCCTTTACAG 360
```

```

Qy 1969 CTATCAAGTCTCTGAAGATGACCCCACTTGCNAAAACAAAGGAATTTTGGCCACCAATTTTG 2028
Db 361 CTATCAAGTCTCTGAAGATGACCCCACTTGCNAAAACAAAGGAATTTTGGCCACCAATTTTG 420

Qy 2029 GATTTCAGCAGGTCTCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTGAAAGTACT 2088
Db 421 GATTTCAGCAGGTCTCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTGAAAGTACT 480

Qy 2089 ACTCGATTTACATGTTATGGGATGCTCTACAAGCCTCATGATCATACGCCCTGGAAAGAAA 2148
Db 481 ACTCGATTTACATGTTATGGGATGCTCTACAAGCCTCATGATCATACGCCCTGGAAAGAAA 540

Qy 2149 TATCCTACTGTGCTGTTTATATATGTTGCTCTCAGGTGCAGTTCGTTGAATATCGGTTT 2208
Db 541 TATCCTACTGTGCTGTTTATATATGTTGCTCTCAGGTGCAGTTCGTTGAATATCGGTTT 600

Qy 2209 AAAGGAGTCAAGTATTTCCGCTTTGAATACCCTAGCCTCTCTAGGTTATGTGGTGTAGTG 2268
Db 601 AAAGGAGTCAAGTATTTCCGCTTTGAATACCCTAGCCTCTCTAGGTTATGTGGTGTAGTG 660

Qy 2269 ATAGACAAACAGGGGATCCTGTCAACCGGGCTTTAAATTTGAAGGCGCTTTAAATATAAAA 2328
Db 661 ATAGACAAACAGGGGATCCTGTCAACCGGGCTTTAAATTTGAAGGCGCTTTAAATATAAAA 720

Qy 2329 ATGGGTCAAAATACAAATTCACGATCAGGTGG 2359
Db 721 ATGGGTCAAAATACAAATTCACGATCAGGTGG 751

RESULT 13
CS998849
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 835)
Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
and Korn,B.
Human T-Lymphocytes library
Unpublished (2005)
Contact: Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9016K2027.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No. 9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4, Primer sequence: CGGATAACAATTCACACAG.
Location/Qualifiers
1..835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016K2027"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
```





```
QY 1032 TGAACAACTCCGAGTGGTGGTAAATCTTAGAATCTATATGAAGAAATGATGAATC 1091
Db 485 TGAACAACTCCGAGTGGTGGTAAATCTTAGAATCTATATGAAGAAATGATGAATC 544
QY 1092 TGAGGTGGAATTAATTCATGTTACATCCCTATGTTGGAACAAGAGGGGAGATTCATT 1151
Db 545 TGAGGTGGAATTAATTCATGTTACATCCCTATGTTGGAACAAGAGGGGAGATTCATT 604
QY 1152 CCGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAGATGTCAGAAATAAT 1211
Db 605 CCGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACTTTTAAGATGTCAGAAATAAT 664
QY 1212 GATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAGGAACCTTAATTCACCTTTTGA 1271
Db 665 GATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAGGAACCTTAATTCACCTTTTGA 724
QY 1272 GATTCTATTTGAAGGAGT 1289
Db 725 GATTCTATTTGAAGGAGT 742

RESULT 15
BX390898
LOCUS BX390898 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL005YD02 5-PRIME, mRNA sequence.
ACCESSION BX390898
VERSION BX390898.2 GI:46844267
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 914)
Li.W.B., Gruber.C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30607432.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7542.x
For More information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAG0532D07_CS05044_1&c=7542.x

FEATURES
Source Location/Qualifiers
1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YD02"
/call_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."

ORIGIN
Query Match 22.0%; Score 686; DB 5; Length 914;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 736; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 739 GGAGGCCACAGGATTTAGCGAAACAACTTTTAAAGCCCAATCTAGTCGAAACTAGTTGT 798
Db 73 GGAGGCCACAGGATTTAGCGAAACAACTTTTAAAGCCCAATCTAGTCGAAACTAGTTGT 132
QY 799 CCCAACATACAGATGGATTCAAAATTTATGCCCGCTGATCCAGACTCGATTTGCTTTTATA 858
Db 133 CCCAACATACAGATGGATTCAAAATTTATGCCCGCTGATCCAGACTCGATTTGCTTTTATA 192
QY 859 CATAGCAACGATTTTGGATATCTAAATCATCGTAAACAGAGAAAGAGAGACTCACTTAT 918
Db 193 CATAGCAACGATTTTGGATATCTAAATCATCGTAAACAGAGAAAGAGAGACTCACTTAT 252
QY 919 GTGCACAAATGAGCTAGCAGCAACATGGAAGAGATGCCAGATCAGCTGGAGTCCGCTACCTTT 978
Db 253 GTGCACAAATGAGCTAGCAGCAACATGGAAGAGATGCCAGATCAGCTGGAGTCCGCTACCTTT 312
QY 979 GTTCTCTCAAGAAGAAATTTGATAGATATCTGGCTATTTGGTGGTCTCCAAAAGCTGAAACA 1038
Db 313 GTTCTCTCAAGAAGAAATTTGATAGATATCTGGCTATTTGGTGGTCTCCAAAAGCTGAAACA 372
QY 1039 ACTCCCAAGTGGTGAATAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTG 1098
Db 373 ACTCCCAAGTGGTGAATAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTG 432
QY 1099 GAAATTTATCATGTTACATCCCTATGTTGGAACAAGAGGGGAGATTCATTCCGTTAT 1158
Db 433 GAAATTTATCATGTTACATCCCTATGTTGGAACAAGAGGGGAGATTCATTCCGTTAT 492
QY 1159 CCTAARACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGTGTCAAGAAATGATGATTTGAT 1218
Db 493 CCTAARACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGTGTCAAGAAATGATGATTTGAT 552
QY 1219 GCTGAAGGAAGGATCATAGATGTCATAGATAGGAACCTTAATTCACCTTTTGAGATTTCTA 1278
Db 553 GCTGAAGGAAGGATCATAGATGTCATAGATAGGAACCTTAATTCACCTTTTGAGATTTCTA 612
QY 1279 TTTGAAGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATATGCTTGG 1338
Db 613 TTTGAAGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATATGCTTGG 672
QY 1339 TCCATCTCTACTAGATCGCTCCCAAGACTCGCTACAGATAGTGTTCATCTCACCTGAATTA 1398
Db 673 TCCATCTCTACTAGATCGCTCCCAAGACTCGCTACAGATAGTGTTCATCTCACCTGAATTA 732
QY 1399 TTTATCCCAAGTGAAGATGATGTTATGGAAGGAGAGAGACTCATTTGAGTGCCTGAT 1458
Db 733 TTTATCCCAAGTGAAGATGATGTTATGGAAGGAGAGAGACTCATTTGAGTGCCTGAT 792
QY 1459 TCTGTGACGCCACTAAT 1475
Db 793 TCTGTGACGCCACTAAT 809

Search completed: May 4, 2006, 12:52:35
Job time : 11802 secs
```



Db 241 CTGGGTGTGAGATATTGTGAACCTGGGACTGTGAGGAGATATTGAATCACAGGATCGG 300  
Qy CTTAAATTTGGAGCCCTTTTATGTATTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360  
Db CTTAAATTTGGAGCCCTTTTATGTATTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360  
Qy GCCGATACCCAGAAAATCATGCTCATGATGCGCTTAGGCCACACACATGATTTTCATGTTT 420  
Db GCCGATACCCAGAAAATCATGCTCATGATGCGCTTAGGCCACACACATGATTTTCATGTTT 420  
Qy GTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGAACTATTACCTTGGCCATGCT 480  
Db GTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGAACTATTACCTTGGCCATGCT 480  
Qy GGTGAGACACAGAGAAATACACTGTTTATTCGAAAATCCCAAACATCAATAGAGCA 540  
Db GGTGAGACACAGAGAAATACACTGTTTATTCGAAAATCCCAAACATCAATAGAGCA 540  
Qy CGAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
Db CGAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
Qy GGAATGPTATCTCGAGAAGAAGAACTATTAAAGAGAAAGAAACGCATTTGGAACAGTCGGA 660  
Db GGAATGPTATCTCGAGAAGAAGAACTATTAAAGAGAAAGAAACGCATTTGGAACAGTCGGA 660  
Qy ATTGCTTCTTACGATTTATCACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA 720  
Db ATTGCTTCTTACGATTTATCACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGA 720  
Qy ATTATCACGTAAAGATGGAGGCCACAAAGGATTTACGCAACACCTTTTAAAGGCCCAT 780  
Db ATTATCACGTAAAGATGGAGGCCACAAAGGATTTACGCAACACCTTTTAAAGGCCCAT 780  
Qy CTAGTGAAGAACTAGTTGTCACACATACGGATGGATCCAAATTTATGCTCCGCTGATCCA 840  
Db CTAGTGAAGAACTAGTTGTCACACATACGGATGGATCCAAATTTATGCTCCGCTGATCCA 840  
Qy GACTGGAATTCCTTTATACATAGCAACCATATTGGATATCTAAATCTGTAACACAGAA 900  
Db GACTGGAATTCCTTTATACATAGCAACCATATTGGATATCTAAATCTGTAACACAGAA 900  
Qy GAAGAGACTCACTTATGTGCAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA 960  
Db GAAGAGACTCACTTATGTGCACATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCA 960  
Qy GCTGGAGTCGCTACCTTTGTTCTCAAAGAAATTTGATAGATATTCTGGCTATTGGTGG 1020  
Db GCTGGAGTCGCTACCTTTGTTCTCAAAGAAATTTGATAGATATTCTGGCTATTGGTGG 1020  
Qy TGCCAAAAGCTGAAACAACTCCAGTGGTGGTAAAAATTTCTAGAAATTTCTATAGAA 1080  
Db TGCCAAAAGCTGAAACAACTCCAGTGGTGGTAAAAATTTCTAGAAATTTCTATAGAA 1080  
Qy AATGATGAATCTGAGGTGAAATTTATTCATGTTACATCCCTATGTTGAAACAAAGGAG 1140  
Db AATGATGAATCTGAGGTGAAATTTATTCATGTTACATCCCTATGTTGAAACAAAGGAG 1140  
Qy GCAGATTCATTCGGTTATCTTAAACAGGTACACAAATCCTAAAGTCACCTTTTAAGATG 1200  
Db GCAGATTCATTCGGTTATCTTAAACAGGTACACAAATCCTTAAAGTCACCTTTTAAGATG 1200  
Qy TCAGAAATTAATGATGCTGMAAGGAGTATAGATCTCATAGATAGAACTAAT 1260  
Db TCAGAAATTAATGATGCTGMAAGGAGTATAGATCTCATAGATAGAACTAAT 1260  
Qy CAACCTTTGAGATCTAATTTGMAAGGAGTGAATATATGCCAGACTGGATGCTCT 1320  
Db CAACCTTTGAGATCTAATTTGMAAGGAGTGAATATATGCCAGACTGGATGCTCT 1320  
Qy GAGGAAAATATGCTTGTGTCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380  
Db GAGGAAAATATGCTTGTGTCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380

Qy 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC 1440  
Db 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC 1440  
Qy 1441 ATTGAGTCAGTGCCTGATTTCTGFGAGCCACATAATTTATCTATGAAGAAACACAGACATC 1500  
Db 1441 ATTGAGTCAGTGCCTGATTTCTGFGAGCCACATAATTTATCTATGAAGAAACACAGACATC 1500  
Qy 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAGTCACGAAGAGGAAATTCAG 1560  
Db 1501 TGSATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAGTCACGAAGAGGAAATTCAG 1560  
Qy 1561 TTTATTTTTCCTCTCGAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTAT 1620  
Db 1561 TTTATTTTTCCTCTCGAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTAT 1620  
Qy 1621 TTAAGGAAAAGCAAAATATAAAGATCCAGTGGTGGCTGCTGCTCCAAAGTATTCAAG 1680  
Db 1621 TTAAGGAAAAGCAAAATATAAAGATCCAGTGGTGGCTGCTGCTCCAAAGTATTCAAG 1680  
Qy 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGAGTTCTTTGGCCGGCAT 1740  
Db 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGGTGAATGGAGTTCTTTGGCCGGCAT 1740  
Qy 1741 GGATCTAATATCCAAAGTTCAGTGCAGAGGCTGGTATATTTTGAAGGCACCAAGAC 1800  
Db 1741 GGATCTAATATCCAAAGTTCAGTGCAGAGGCTGGTATATTTTGAAGGCACCAAGAC 1800  
Qy 1801 TCCGCTTTAGAGCATCACCTGTACGTAGTCACTAGTAAATCTCGAGAGGTGACAAG 1860  
Db 1801 TCCGCTTTAGAGCATCACCTGTACGTAGTCACTAGTAAATCTCGAGAGGTGACAAG 1860  
Qy 1861 CTGACTGACCGTGGCTACTCAATTTCTGCTGCACTAGTCAGCACTGTGACTTCTTTATA 1920  
Db 1861 CTGACTGACCGTGGCTACTCAATTTCTGCTGCACTAGTCAGCACTGTGACTTCTTTATA 1920  
Qy 1921 AGTAAGTATAGTAAACAGAGAAATCCACATCTGTGTGCTCCCTTTTCAAGACTATCAAGTCCT 1980  
Db 1921 AGTAAGTATAGTAAACAGAGAAATCCACATCTGTGTGCTCCCTTTTCAAGACTATCAAGTCCT 1980  
Qy 1981 GAAGATGACCCAACTTGGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040  
Db 1981 GAAGATGACCCAACTTGGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040  
Qy 2041 CCTCTTCTGACTATACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTTACA 2100  
Db 2041 CCTCTTCTGACTATACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTTACA 2100  
Qy 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCTCGAAGAAATATCCTACTGTG 2160  
Db 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCTCGAAGAAATATCCTACTGTG 2160  
Qy 2161 CTGTTTCATATATGTTGTTCTCAGGTGCGATTTGGTGAATTAATCGGTTTAAAGAGTCAAG 2220  
Db 2161 CTGTTTCATATATGTTGTTCTCAGGTGCGATTTGGTGAATTAATCGGTTTAAAGAGTCAAG 2220  
Qy 2221 TATTTCCGCTTGAATACCTTAGCTCTCTAGGTTATGTTGTTAGTATAGACACAGG 2280  
Db 2221 TATTTCCGCTTGAATACCTTAGCTCTCTAGGTTATGTTGTTAGTATAGACACAGG 2280  
Qy 2281 GGATCTCTGFCACCGAGGCTTAAATTTGAAGGCCCTTTTAAATATATAAATTTGGTCAAATA 2340  
Db 2281 GGATCTCTGFCACCGAGGCTTAAATTTGAAGGCCCTTTTAAATATATAAATTTGGTCAAATA 2340  
Qy 2341 GAAATTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGATTTTCAATGAC 2400  
Db 2341 GAAATTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGATTTTCAATGAC 2400  
Qy 2401 TTAGATCTGTGGGCATCCAGCGTGTGCTTATGGAGGATACCTCTCCCTGTAGGCAATTA 2460  
Db 2401 TTAGATCTGTGGGCATCCAGCGTGTGCTTATGGAGGATACCTCTCCCTGTAGGCAATTA 2460

QY 2461 ATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC 2520  
DB 2461 ATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC 2520  
QY 2521 TTCTATGATACAGGATACAGGAACGTTATATGGTTCACCCCTGACCAAGATCAACAGGC 2580  
DB 2521 TTCTATGATACAGGATACAGGAACGTTATATGGTTCACCCCTGACCAAGATCAACAGGC 2580  
QY 2581 TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640  
DB 2581 TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640  
QY 2641 CTGCTCTTACATGTTTCTGGATGAGAAATGTCCTATTTTGCATACACAGATATATTCTG 2700  
DB 2641 CTGCTCTTACATGTTTCTGGATGAGAAATGTCCTATTTTGCATACACAGATATATTCTG 2700  
QY 2701 AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACAC 2760  
DB 2701 AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACAC 2760  
QY 2761 AGCATAAGAGTTCTGAAATCGGAGAACATTAATGAACTGCACTCTTTTGCATCACTTCAA 2820  
DB 2761 AGCATAAGAGTTCTGAAATCGGAGAACATTAATGAACTGCACTCTTTTGCATCACTTCAA 2820  
QY 2821 GAAAACCTTGGATCAGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA 2940  
DB 2821 GAAAACCTTGGATCAGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGA 2940  
QY 2941 ATTGATCATCATCTTTGATACCTGTCATGTAACATCTACTCTGAAAATAAATGTGGTG 3000  
DB 2941 ATTGATCATCATCTTTGATACCTGTCATGTAACATCTACTCTGAAAATAAATGTGGTG 3000  
QY 3001 CcATGCAGGGCTACGGTTTGTGGTAGTAATCTAACTTAAACCCCAATGCTCAAAA 3060  
DB 3001 CcATGCAGGGCTACGGTTTGTGGTAGTAATCTAACTTAAACCCCAATGCTCAAAA 3060  
QY 3061 TCAATGATACATATCTCTGAGAGACCCAGCAATCAATGATTAATCAATGATTAATCAAAA 3120  
DB 3061 TCAATGATACATATCTCTGAGAGACCCAGCAATCAATGATTAATCAATGATTAATCAAAA 3120

RESULT 2

US-09-976-674-12  
; Sequence 12, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976, 674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 4829  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-12

Query Match 92.6%; Score 2889; DB 3; Length 4829;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;  
QY 1 AAGTGTAAAGCCTCCGAGGCCAAGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGGCG 60

DB 1 AAGTGTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGGCG 60  
QY 61 CGTTCCGCGCTGGGTTGTACCGCGCCCGCGCGGGAAGCACTGCAACACAGGACCG 120  
DB 61 CGTTCCGCGCTGGGTTGTACCGCGCCCGCGCGGGAAGCACTGCAACACAGGACCG 120  
QY 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGAGCGG 180  
DB 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGAGCGG 180  
QY 181 TCCGGCGCGCGCGGGAAGAAAATGCAACATGCGCAGCAGCAATGGAACACAGAACAG 240  
DB 181 TCCGGCGCGCGCGGGAAGAAAATGCAACATGCGCAGCAGCAATGGAACACAGAACAG 240  
QY 241 CTGGGTGTTGAGATATTTGAAAATGCGGCACTGTGAGGAGAAATTTGAATACAGGATCGG 300  
DB 241 CTGGGTGTTGAGATATTTGAAAATGCGGCACTGTGAGGAGAAATTTGAATACAGGATCGG 300  
QY 301 CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAAAGCTGCTT 360  
DB 301 CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAAAGCTGCTT 360  
QY 361 GCCGATACAGAAAATATCATGGCTACATGCTAAGGCAACCATGATGATTTTCAATGTTT 420  
DB 361 GCCGATACAGAAAATATCATGGCTACATGCTAAGGCAACCATGATGATTTTCAATGTTT 420  
QY 421 GTGAAGAGGAATGATCCAGATGGAACCTTATTGAGAGCAATCTATTACCTTGGCATGCT 480  
DB 421 GTGAAGAGGAATGATCCAGATGGAACCTTATTGAGAGCAATCTATTACCTTGGCATGCT 480  
QY 481 GGTGAGAAACAGAGAAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA 540  
DB 481 GGTGAGAAACAGAGAAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA 540  
QY 541 GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
DB 541 GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
QY 601 GGAATGTAATCTCGAGAGAAAGAACTATTAAAGAGAAAGAAACGCATTTGGAACAGTCGA 660  
DB 601 GGAATGTAATCTCGAGAGAAAGAACTATTAAAGAGAAAGAAAGCACTTTGGAACAGTCGA 660  
QY 661 ATTGCTTCTTACGATATATCACCAGGAAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGA 720  
DB 661 ATTGCTTCTTACGATATATCACCAGGAAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGA 720  
QY 721 ATTTATCACGTAAGAGATGGAGGCCCAAGGATTTTACGCAACAACTTTTAAAGGCCCAAT 780  
DB 721 ATTTATCACGTAAGAGATGGAGGCCCAAGGATTTTACGCAACAACTTTTAAAGGCCCAAT 780  
QY 781 CTAGTGGAAACTAGTGTCTCCCAACATACAGGATTTTACGCAACAACTTTTAAAGGCCCAAT 840  
DB 781 CTAGTGGAAACTAGTGTCTCCCAACATACAGGATTTTACGCAACAACTTTTAAAGGCCCAAT 840  
QY 841 GACTGGATTTGCTTTTATACATAGCAACAGATATTGGAATCTTAAACATCTGTAACAGAGAA 900  
DB 841 GACTGGATTTGCTTTTATACATAGCAACAGATATTGGAATCTTAAACATCTGTAACAGAGAA 900  
QY 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960  
DB 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960  
QY 961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020  
DB 961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020  
QY 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080  
DB 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA 1080  
QY 1081 AATGATGAATCTGAGGTGGAAATTTTCTGTTTACATCCCTATGTTGGAAACAAAGGAGG 1140

1081	AA	TGATGAATCTGAGGTGGGAATTTATTCATGTTACATCCCTATGTTGGAACAAGGAGG	1140
1141	QY	GGAGATTCATTTCCGTTATCTTAAACAAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGATG	1200
	DB	GGAGATTCATTTCCGTTATCTTAAACAAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGATG	1200
	DB	GGAGATTCATTTCCGTTATCTTAAACAAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGATG	1200
1201	QY	TCAGAAATAATGATTTGATGCTGAAGGAAGGATCATAGATGTCTATAGATAAGGAACCTAATT	1260
1201	DB	TCAGAAATAATGATTTGATGCTGAAGGAAGGATCATAGATGTCTATAGATAAGGAACCTAATT	1260
1261	QY	CAACCTTTTGAGATTCATTTTGAGGAGTTTGAATATATTTGCCAGAGCTTGGATGGACTCCT	1320
1261	DB	CAACCTTTTGAGATTCATTTTGAGGAGTTTGAATATATTTGCCAGAGCTTGGATGGACTCCT	1320
1321	QY	GAGGGAAAATATGCTTTGGTCCATCCTACTAGATCGCTCCACAGCTCGCTACACAGATAGTG	1380
1321	DB	GAGGGAAAATATGCTTTGGTCCATCCTACTAGATCGCTCCACAGCTCGCTACACAGATAGTG	1380
1381	QY	TTGATCTCACCTGGAATTAATTTATCCCACTAGTAGAGATGATGTTATCGAAAGGCACAGACTC	1440
1381	DB	TTGATCTCACCTGGAATTAATTTATCCCACTAGTAGAGATGATGTTATCGAAAGGCACAGACTC	1440
1441	QY	ATTGAGTCAGTGGCTGATCTGTGAGCCACCTAATTAATCTATGAGRAACACACAGATC	1500
1441	DB	ATTGAGTCAGTGGCTGATCTGTGAGCCACCTAATTAATCTATGAGRAACACACAGATC	1500
1501	QY	TGGATAATATCCATGACATCTTTCATGTTTTTCCCRAAGTCACGAGAGAAATTGAG	1560
1501	DB	TGGATAATATCCATGACATCTTTCATGTTTTTCCCRAAGTCACGAGAGAAATTGAG	1560
1561	QY	TTTTATTTTGGCTCTGAAATGCAACAAGTTTCCGTCATTTTATACAAAATTAACATCTAATT	1620
1561	DB	TTTTATTTTGGCTCTGAAATGCAACAAGTTTCCGTCATTTTATACAAAATTAACATCTAATT	1620
1621	QY	TTTAAAGGAAAGCAATATAAACAAGTCCAGTGGTGGCTGCCTGCTCCAAGTGATTTCAAG	1680
1621	DB	TTTAAAGGAAAGCAATATAAACAAGTCCAGTGGTGGCTGCCTGCTCCAAGTGATTTCAAG	1680
1681	QY	TGTCCTATCAAAGGAGAGTAGCAATTAACCAAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT	1740
1681	DB	TGTCCTATCAAAGGAGAGTAGCAATTAACCAAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT	1740
1741	QY	GGATCTAATATCCAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
1741	DB	GGATCTAATATCCAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
1801	QY	TCCCTTTAGAGCATCACTGTAGTAGTCAGTTACGTTAAATCTGTGAGAGGTCACAAAGG	1860
1801	DB	TCCCTTTAGAGCATCACTGTAGTAGTCAGTTACGTTAAATCTGTGAGAGGTCACAAAGG	1860
1861	QY	CTGACTGACCGTGCTACTACATCTTCTGTCGATCAGTCAGCACTGTGACTCTCTTTATA	1920
1861	DB	CTGACTGACCGTGCTACTACATCTTCTGTCGATCAGTCAGCACTGTGACTCTCTTTATA	1920
1921	QY	AGTAGATATAGTAAACCAAGAAATCCACACTGTGTGCTCCTTTACAAGCTATCAAGTCTCT	1980
1921	DB	AGTAGATATAGTAAACCAAGAAATCCACACTGTGTGCTCCTTTACAAGCTATCAAGTCTCT	1980
1981	QY	GAAGATGACCCAACTTCGAAAAACAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
1981	DB	GAAGATGACCCAACTTCGAAAAACAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
2041	QY	CCTCTTCTGACTACTCCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
2041	DB	CCTCTTCTGACTACTCCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
2101	QY	TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTTACTGTG	2160
2101	DB	TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTTACTGTG	2160
2161	QY	CTGTTTCATATATGGTGG--TCCCTCAGGTGCAGTTGGTGAATATCGGTTTAAAGGAGTCA	2218
2161	DB	CTGTTTCATATATGGTGGTCTCTCAGGTGCAGTTGGTGAATATCGGTTTAAAGGAGTCA	2220

### RESULTS

US-09-976-674-2  
; Sequence 2, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen



; APPLICANT: Riviere, Pierre									
; APPLICANT: Junien, Jean-Louis									
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV									
; FILE REFERENCE: 70669									
; CURRENT APPLICATION NUMBER: US/09/976,674									
; CURRENT FILING DATE: 2001-10-12									
; PRIOR APPLICATION NUMBER: US 60/240,117									
; PRIOR FILING DATE: 2000-10-12									
; NUMBER OF SEQ ID NOS: 61									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 2									
; LENGTH: 2671									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-976-674-2									
Query Match 80.0%; Score 2495; DB 3; Length 2671;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;									
QY	213	CATGGCAGCAGCAATGGAACAGAACAGCTGGGTGGAGATATTTGAAACTCGGACTG	272						
DB	7	CATGGCAGCAGCAATGGAACAGAACAGCTGGGTGGAGATATTTGAAACTCGGACTG	66						
QY	273	TGAGGAGAATATTGAATCAAGGATCGGCTTAAATTGGAGCCTTTTATGTTGAGCGGTA	332						
DB	67	TGAGGAGAATATTGAATCAAGGATCGGCTTAAATTGGAGCCTTTTATGTTGAGCGGTA	126						
QY	333	TTCTGAGTCAGCTTAAAGCTGCTGGCCGATACGAGAAATATCATGCTACATGAT	392						
DB	127	TTCTGAGTCAGCTTAAAGCTGCTGGCCGATACGAGAAATATCATGCTACATGAT	186						
QY	393	GGCTAAGGCACCATGATTTTATGTTGTGAGAGGATCATCCAGATGGACTCATTC	452						
DB	187	GGCTAAGGCACCATGATTTTATGTTGTGAGAGGATCATCCAGATGGACTCATTC	246						
QY	453	AGACAGAACTTATTCCTTGCATCTGCTGGTGAGAACAGAGAAATACACTGTTTATTC	512						
DB	247	AGACAGAACTTATTCCTTGCATCTGCTGGTGAGAACAGAGAAATACACTGTTTATTC	306						
QY	513	TGAAATTCCTCAAACTATCAATAGACAGCAGCTTAAATGCTCTCTTGGAGCCTCTTTT	572						
DB	307	TGAAATTCCTCAAACTATCAATAGACAGCAGCTTAAATGCTCTCTTGGAGCCTCTTTT	366						
QY	573	GGATCTTTTTCAGGCAACACTGGAATATGGAATGATTTCTCGAGAAGAACTATTAAAG	632						
DB	367	GGATCTTTTTCAGGCAACACTGGAATATGGAATGATTTCTCGAGAAGAACTATTAAAG	426						
QY	633	AGAAAGAAACCGATTTGGAAACAGTGGAAATTCCTTACGATTTATCACCAGAGAGTGG	692						
DB	427	AGAAAGAAACCGATTTGGAAACAGTGGAAATTCCTTACGATTTATCACCAGAGAGTGG	486						
QY	693	AACATTTCTTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCCACAAAG	752						
DB	487	AACATTTCTTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCCACAAAG	546						
QY	753	ATTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAATCTAGTGTCCCAACATACGGAT	812						
DB	547	ATTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAATCTAGTGTCCCAACATACGGAT	606						
QY	813	GGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATAT	872						
DB	607	GGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATAT	666						
QY	873	TTGGATATCTAACATCGTAAACAGAGAGAAAGGAGACTCTATTATGTCACAATGAGCT	932						
DB	667	TTGGATATCTAACATCGTAAACAGAGAGAAAGGAGACTCTATTATGTCACAATGAGCT	726						
QY	933	AGCCAACTAGGAGAGATGCCAGATCAGCTGGATCGTACCTTTGTTCTCCAGAGA	992						
DB	727	AGCCAACTAGGAGAGATGCCAGATCAGCTGGATCGTACCTTTGTTCTCCAGAGA	786						
QY	993	ATTTGATAGATATTTCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCCGAGTGGTGG	1052						

DB	787	ATTGTAGATATTTCTGGCTATTTGGTGGTGTCCAAAGCTGAAACAACTCCCGTGGTG	846						
QY	1053	TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTCAGGTGGAATTTATTCATGT	1112						
DB	847	TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTCAGGTGGAATTTATTCATGT	906						
QY	1113	TACATCCCTATGTTGGAAACAAAGAGGCGCAGATTCATTCGGTTATCTTAAACAGGTAC	1172						
DB	907	TACATCCCTATGTTGGAAACAAAGAGGCGCAGATTCATTCGGTTATCTTAAACAGGTAC	966						
QY	1173	AGCAAAATCTTAAAGTCACCTTTTAAAGATGTCAAGAAATAATGATGATGCTGAAGAAAGAT	1232						
DB	967	AGCAAAATCTTAAAGTCACCTTTTAAAGATGTCAAGAAATAATGATGATGCTGAAGAAAGAT	1026						
QY	1233	CATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTTCTATTTGAAGAGATTGA	1292						
DB	1027	CATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTTCTATTTGAAGAGATTGA	1086						
QY	1293	ATATATTCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCTACTAGA	1352						
DB	1087	ATATATTCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCTACTAGA	1146						
QY	1353	TCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCACTAGA	1412						
DB	1147	TCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCACTAGA	1206						
QY	1413	AGATGATGTTTATGGAAGGCGAGACTCAATTGAGTCAGTGGCTGATTTCTGTGAGCCACT	1472						
DB	1207	AGATGATGTTTATGGAAGGCGAGACTCAATTGAGTCAGTGGCTGATTTCTGTGAGCCACT	1266						
QY	1473	AAATTTCTATGGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1532						
DB	1267	AAATTTCTATGGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1326						
QY	1533	TCGCCAAAGTCAGAAAGAGAAATTTGAGTTTATTTTGGCTCTGATGATCCAAAACAGGTTT	1592						
DB	1327	TCGCCAAAGTCAGAAAGAGAAATTTGAGTTTATTTTGGCTCTGATGATCCAAAACAGGTTT	1386						
QY	1593	CCGTCAATTTATCAAAATTTACATCTATTTTAAAGAAACGAAATATAACCATCCAGTGG	1652						
DB	1387	CCGTCAATTTATCAAAATTTACATCTATTTTAAAGAAACGAAATATAACCATCCAGTGG	1446						
QY	1653	TGGCTCCCTGCTCCAAAGTGTCTTCAAGTGCTCTTCAAGAGGAGATAGCAATTTACCAG	1712						
DB	1447	TGGCTCCCTGCTCCAAAGTGTCTTCAAGTGCTCTTCAAGAGGAGATAGCAATTTACCAG	1506						
QY	1713	TGGTGAATGGGAAGTTCTTTGGCCGCGCATGGAATCTAATATCCAAAGTTGATGAAGTCAGAAG	1772						
DB	1507	TGGTGAATGGGAAGTTCTTTGGCCGCGCATGGAATCTAATATCCAAAGTTGATGAAGTCAGAAG	1566						
QY	1773	GCTGGTATATTTTGAAGGCACCAAGACCTCCCTTTTGAAGCATACCTGTACGTAGTCAG	1832						
DB	1567	GCTGGTATATTTTGAAGGCACCAAGACCTCCCTTTTGAAGCATACCTGTACGTAGTCAG	1626						
QY	1833	TTACGTAATCTTGGAGAGGTGACAAAGGTGACTGACCGTGGCTACTCACATTTCTGCTG	1892						
DB	1627	TTACGTAATCTTGGAGAGGTGACAAAGGTGACTGACCGTGGCTACTCACATTTCTGCTG	1686						
QY	1893	CATCAGTCAGCATCTGTGACTTCTTTTATAGTAAGTATAGTACACGAGAGAAATCCACACTG	1952						
DB	1687	CATCAGTCAGCATCTGTGACTTCTTTTATAGTAAGTATAGTACACGAGAGAAATCCACACTG	1746						
QY	1953	TGTCGCTCTTTTACAAAGCTATCAAGTCTCTGAAAGTACCCAACTTGCACAAAACAAAGGAAT	2012						
DB	1747	TGTCGCTCTTTTACAAAGCTATCAAGTCTCTGAAAGTACCCAACTTGCACAAAACAAAGGAAT	1806						
QY	2013	TTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTCGTACTATCTCTCCAGAAATTTT	2072						
DB	1807	TTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTCGTACTATCTCTCCAGAAATTTT	1866						
QY	2073	CTCTTTTGAAGATCTACTGGAATTTACATGTAATGGATGCTCTTCAAGCCCTCATGATCT	2132						

Db 1867 CTCTTTTAAAGTACTTGGATTTTACATTTGTATGGATGCTCTACAGGCTCATGATCT 1926  
Qy 2133 ACAGCCTGGAAAGAAATATCCTACTGCTGTTTCATATATGTTGGTCTCTCAGGTGAGTT 2192  
Db 1927 ACAGCCTGGAAAGAAATATCCTACTGCTGTTTCATATATGTTGGTCTCTCAGGTGAGTT 1986  
Qy 2193 GGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCAGG 2252  
Db 1987 GGTGAATAATTCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCAGG 2046  
Qy 2253 TTATGTTGTTGTTAGTATGATACACACAGGGGATCTGTCAACGAGGGCTTTAAATTTGAAG 2312  
Db 2047 TTATGTTGTTGTTAGTATGATACACACAGGGGATCTGTCAACGAGGGCTTTAAATTTGAAG 2106  
Qy 2313 CGCCTTTAAATATAAATGGGTCAATATAGAAATGACGATCAGGTGGAAGGACTCCAAATA 2372  
Db 2107 CGCCTTTAAATATAAATGGGTCAATATAGAAATGACGATCAGGTGGAAGGACTCCAAATA 2166  
Qy 2373 TCTAGCTTCTCGATATGATTTCAATTTAGCTTATGATCGTGTGGCATCCACGGCTGGTCTTA 2432  
Db 2167 TCTAGCTTCTCGATATGATTTCAATTTAGCTTATGATCGTGTGGCATCCACGGCTGGTCTTA 2226  
Qy 2433 TGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTTCAGATATCTTTCAGGGTTGCTAT 2492  
Db 2227 TGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTTCAGATATCTTTCAGGGTTGCTAT 2286  
Qy 2493 TGTCTGGGGCCCCAGTCACTCTCTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 2552  
Db 2287 TGTCTGGGGCCCCAGTCACTCTCTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 2346  
Qy 2553 GGGTCACTCCGACCAAGATGACAGGGCTATTTACTGCTCTTACATGCTTTCTCGATGAGAAATGT 2672  
Db 2347 GGGTCACTCCGACCAAGATGACAGGGCTATTTACTGCTCTTACATGCTTTCTCGATGAGAAATGT 2466  
Qy 2673 CCATTTTGACATACACAGTATATTTACTGAGTTTTTTTATGAGGGCTGGAAGCCATATGA 2732  
Db 2467 CCATTTTGACATACACAGTATATTTACTGAGTTTTTTTATGAGGGCTGGAAGCCATATGA 2526  
Qy 2733 TTATACAGATCTATCTCTAGGAGACACAGCATTAAGAGTTTCTGAAATCGGAGAACATTA 2792  
Db 2527 TTATACAGATCTATCTCTAGGAGACACAGCATTAAGAGTTTCTGAAATCGGAGAACATTA 2586  
Qy 2793 TGAATCGCATCTTTTGCACTACCTTTCAAGAAAACCTTTGGATCAGTATTTGCTCTTAA 2852  
Db 2587 TGAATCGCATCTTTTGCACTACCTTTCAAGAAAACCTTTGGATCAGTATTTGCTCTTAA 2646  
Qy 2853 AGTGATAT 2860  
Db 2647 AGTGATAT 2654

RESULT 4  
US-09-976-594-1103  
; Sequence 1103, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Fumess, Michael  
; APPLICANT: Buchinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 1103  
; LENGTH: 2797  
; TYPE: DNA

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1  
US-09-976-594-1103

Query Match 74.2%; Score 2315; DB 3; Length 2797;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;

Qy 335 CCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACACAGAAAAATATCATGGCTACATGATGG 394  
Db 1 CCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACACAGAAAAATATCATGGCTACATGATGG 60  
Qy 395 CTAAGGCAACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAG 454  
Db 61 CTAAGGCAACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAG 120  
Qy 455 ACAGAAATCTATTACCTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAG 514  
Db 121 ACAGAAATCTATTACCTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAG 180  
Qy 515 AAATTTCCCAAACTCAATATAGAGCAGCAGCTTAAATGCTCTCTTGGAGCCTCTTTGG 574  
Db 181 AAATTTCCCAAACTCAATATAGAGCAGCAGCTTAAATGCTCTCTTGGAGCCTCTTTGG 240  
Qy 575 ATCTTTTTCAGGCAACATCGACTATGGAATGTATTTCTCGAGAAAGAACTATTAAGAG 634  
Db 241 ATCTTTTTCAGGCAACATCGACTATGGAATGTATTTCTCGAGAAAGAACTATTAAGAG 300  
Qy 635 AAAGAAAACGCAATGGAACAGTCGGAATTTGCTTTTACGATTTATCACCAAGGAAGTGAA 694  
Db 301 AAAGAAAACGCAATGGAACAGTCGGAATTTGCTTTTACGATTTATCACCAAGGAAGTGAA 360  
Qy 695 CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTAAGAGATGGAGGGCCCAAGGAT 754  
Db 361 CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTAAGAGATGGAGGGCCCAAGGAT 420  
Qy 755 TTAACGCAACCACTTTAAAGGCCCAATCTAGTGGAAACTAGTTTGTCCCAACATACGATATTT 874  
Db 421 TTAACGCAACCACTTTAAAGGCCCAATCTAGTGGAAACTAGTTTGTCCCAACATACGATATTT 480  
Qy 815 ATCCAAAAATATGCCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTT 874  
Db 481 ATCCAAAAATATGCCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATTT 540  
Qy 875 GGATATCTAAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCAACATGAGCTAG 934  
Db 541 GGATATCTAAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCAACATGAGCTAG 600  
Qy 935 CCACATGGAAGAGATGCGAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAAAT 994  
Db 601 CCACATGGAAGAGATGCGAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAAAT 660  
Qy 995 TTGATAGATATTTCTGGCTATTTGGTGTCCAAAAGCTGAAACAACTCCAGGTGGTGA 1054  
Db 661 TTGATAGATATTTCTGGCTATTTGGTGTCCAAAAGCTGAAACAACTCCAGGTGGTGA 720  
Qy 1055 AAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTA 1114  
Db 721 AAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTA 780  
Qy 1115 CATCCCTATGTTGGAACAAAGG-AGGCGAGATTCATTCCGTTTATCCTAAAACAGGTACA 1173  
Db 781 CATCCCTATGTTGGAACAAAGGCGAGGCGAGATTCATTCCGTTTATCCTAAAACAGGTACA 840  
Qy 1174 GCAAAATCCTAAAGTCACCTTTTAAAGATGTCAAGAAATATGATGCTGAAGGAAGGATC 1233  
Db 841 GCAAAATCCTAAAGTCACCTTTTAAAGATGTCAAGAAATATGATGCTGAAGGAAGGATC 900  
Qy 1234 ATAGATGTATAGATAAGGAACTAAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAA 1293  
Db 901 ATAGATGTATAGATAAGGAACTAAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTTGAA 960

QY	1294	TATATTGGCAGAGCTGGATGGACTCTCTGAGGGAAAATA-----TGCTTGGTCCATCTCTACT	1349
Db	961	TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAATAGTGTCTTGGTCCATCTCTACT	1020
QY	1350	AGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCACCCTGAAATTTATTTATCCAGT	1409
Db	1021	AGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCACCCTGAAATTTATTTATCCOAGT	1080
QY	1410	AGAAGATGATCTTATGTGAAAGGCAGAGACTCAATTCAGTTCAGTGTGCTGATTTCTGTGACGCC	1469
Db	1081	AGAAGATGATCTTATGTGAAAGGCAGAGACTCAATTCAGTTCAGTGTGCTGATTTCTGTGACGCC	1140
QY	1470	ACTAATTATCTATCAAGAAAACACAGACATCTGGATAAATATCCATGACATCTTTTCAATGT	1529
Db	1141	ACTAATTATCTATCAAGAAAACACAGACATCTGGATAAATATCCATGACATCTTTTCAATGT	1200
QY	1530	TTTTTCCCAAGTCAAGAAGGAAATTCAGTTTATTTTTTGCCTCTCGAATGCAAAACAGG	1589
Db	1201	TTTTTCCCAAGTCAAGAAGGAAATTCAGTTTATTTTTTGCCTCTCGAATGCAAAACAGG	1260
QY	1590	TTTTCCGTCAATTTATACAAAATATACATCTATTTTTTAAAGGAAAGCAAAATATAAAGATCCAG	1649
Db	1261	TTTTCCGTCAATTTATACAAAATATACATCTATTTTTTAAAGGAAAGCAAAATATAAAGATCCAG	1320
QY	1650	TGAGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAAGAGGAGATAGCAATTAC	1709
Db	1321	TGAGTGGCTGCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAAGAGGAGATAGCAATTAC	1380
QY	1710	CAGTGGTCAATGGGAAGTCTTTCGGCGGCATGATCTAATATCCAAGTTGATGAAGTCAG	1769
Db	1381	CAGTGGTGAATGGGAAGTCTTTCGGCGGCATGATCTAATATCCAAGTTGATGAAGTCAG	1440
QY	1770	AAGGCTGGTATATTTTGAAGGCACCAAAGACTCCCTTTTAGAGCATCACTGTACGTAGT	1829
Db	1441	AAGGCTGGTATATTTTGAAGGCACCAAAGACTCCCTTTTAGAGCATCACTGTACGTAGT	1500
QY	1830	CAGTTACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTTCACATTTCTTG	1889
Db	1501	CAGTTACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTTCACATTTCTTG	1560
QY	1890	CTGCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAGAAATCCACA	1949
Db	1561	CTGCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAGAAATCCACA	1620
QY	1950	CTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGCAAAAACAAAGGA	2009
Db	1621	CTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGCAAAAACAAAGGA	1680
QY	2010	ATTTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTCTGACTACTACTCTCTCCAGAAAT	2069
Db	1681	ATTTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTCTGACTACTACTCTCTCCAGAAAT	1740
QY	2070	TTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTCAAGAGCTCTATGA	2129
Db	1741	TTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTCAAGAGCTCTATGA	1800
QY	2130	TCTACAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGTGTGCTCTCAGGTGCA	2189
Db	1801	TCTACAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGTGTGCTCTCAGGTGCA	1860
QY	2190	GTTTGGTGAATATCGTTTAAAGCAGTCAAGTATTTCCGCTTTGAATACCTAGAGCTCTCT	2249
Db	1861	GTTTGGTGAATATCGTTTAAAGCAGTCAAGTATTTCCGCTTTGAATACCTAGAGCTCTCT	1920
QY	2250	AGGTTATGTGGTTGTAGTGATAGACAAAGGGGATCTCTGTCAACGAGGGGCTTAAATTGA	2309
Db	1921	AGGTTATGTGGTTGTAGTGATAGACAAAGGGGATCTCTGTCAACGAGGGGCTTAAATTGA	1980
QY	2310	AGGCGCTTTAAATATAAATATGAATTTGACGATTCAGTGTGAGGAGCTTCCA	2369
Db	1981	AGGCGCTTTAAATATAAATATGAATTTGACGATTCAGTGTGAGGAGCTTCCA	2040

RESULT 5  
US-09-976-674-8  
; Sequence 8, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976.674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 4523  
; TYPE: DNA

; ORGANISM: Homo sapiens  
US-09-976-674-8

Query Match 68.2%; Score 2128; DB 3; Length 4523;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2178; Conservative 0; Mismatches 1; Indels 0

Qy	1	AAGTGTAAAGCCTTCGAGGCCCAAGCCCGCTGCTACTATGCGCCGCTGCTTCTTTAGTGGCG	60
Db	1	AAGTGTAAAGCCTTCGAGGCCCAAGCCCGCTGCTACTATGCGCCGCTGCTTCTTTAGTGGCG	60
Qy	61	CGTTCGCCGCTGGGTTGTCA CCGGCGCCGCGCGAGGAAGCCACTGCAAC CAGACCG	120
Db	61	CGTTCGCCGCTGGGTTGTCA CCGGCGCCGCGCGAGGAAGCCACTGCAAC CAGACCG	120
Qy	121	GAGTGGAGCGGCGCAGCATGAAGCGGCGCAGGCCCGCTCATAGCGCACGTGGGACGG	180
Db	121	GAGTGGAGCGGCGCAGCATGAAGCGGCGCAGGCCCGCTCATAGCGCACGTGGGACGG	180
Qy	181	TCCGGCGGGCGCGGGGAAGGAAAATGCAACAATGCGCAGCAGCAATGGAAC CAGAACAG	240
Db	181	TCCGGCGGGCGCGGGGAAGGAAAATGCAACAATGCGCAGCAGCAATGGAAC CAGAACAG	240
Qy	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGAGAAATATTGAATCACAAGATCGG	300
Db	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGAGAAATATTGAATCACAAGATCGG	300
Qy	301	CCTAAATTGAGCCCTTTTATGTTGAGCGGTATTCTCTGAGTCAAGTCTTAAAGCTGCTT	360
Db	301	CCTAAATTGAGCCCTTTTATGTTGAGCGGTATTCTCTGAGTCAAGTCTTAAAGCTGCTT	360
Qy	361	GCCGATACCAAGAAAATATCATGGCTACATGATGCTTAAGGCACCAATGATTTTCATGTTT	420
Db	361	GCCGATACCAAGAAAATATCATGGCTACATGATGCTTAAGGCACCAATGATTTTCATGTTT	420
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTTGAAATTTCCCAAACTATCAATAGACGA	480
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCATTTGAAATTTCCCAAACTATCAATAGACGA	480
Qy	481	GGTGAGAACAGAGAAAATACACTGTTTTATTCTGAAATTTCCCAAACTATCAATAGACGA	540
Db	481	GGTGAGAACAGAGAAAATACACTGTTTTATTCTGAAATTTCCCAAACTATCAATAGACGA	540
Qy	541	GCAGTCTTAATGCTCTTTGGAAGCCCTTTTGGATCTTTTCAGGACACCTGACATAT	600
Db	541	GCAGTCTTAATGCTCTTTGGAAGCCCTTTTGGATCTTTTCAGGACACCTGACATAT	600
Qy	601	GGAAATGATCTTCGAGAGAGACTATTAAAGGAAGAAGAACGCAATGGAACAGCTCGGA	660
Db	601	GGAAATGATCTTCGAGAGAGACTATTAAAGGAAGAAGAACGCAATGGAACAGCTCGGA	660
Qy	661	ATTGCTCTTACGATATCACCRAGGAAGCTGGAACATTTCTGTTTCAAGCCGTATGTGA	720
Db	661	ATTGCTCTTACGATATCACCRAGGAAGCTGGAACATTTCTGTTTCAAGCCGTATGTGA	720
Qy	721	ATTTATCACGTTAAGAATGGAGGCCCAAGAGATTTACGCAACACTTTAAGGCCCAAT	780
Db	721	ATTTATCACGTTAAGAATGGAGGCCCAAGAGATTTACGCAACACTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTTGTGCCACATACCGATGGATCCAAAATTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAAACTAGTTGTGCCACATACCGATGGATCCAAAATTATGCCCGCTGATCCA	840
Qy	841	GACTCGATTGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAC CAGAA	900
Db	841	GACTCGATTGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAC CAGAA	900
Qy	901	GAAGAGGACATCACTTATGTGCACAATAGAGCTAGCCCAATGGAAGAGAGATGCCAGATCA	960
Db	901	GAAGAGGACATCACTTATGTGCACAATAGAGCTAGCCCAATGGAAGAGAGATGCCAGATCA	960
Qy	961	GCTGAGTGCCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGTGG	1020

	2101	2101	2101	2160
Qy	TGTTATGGATGCTCTCAAGCCCTCATGATCTACGCTTGGAAAGAAATATCCTACTGTG			2160
Db	TGTTATGGATGCTCTCAAGCCCTCATGATCTACGCTTGGAAAGAAATATCCTACTGTG			2160
Qy	CTGTTCAVATATGTGGTGC	2179		
Db	CTGTTCAVATATGTGGTGC	2179		

## RESULT 6

```

US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70869
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20

```

Query Match	68.2%;	Score 2128;	DB 3;	Length 4676;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 2178;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	AAGTCTAAAGCTTCGAGGCGCAAGGCGCTGCTACTGCGGCGGCTGCTTCTTAGTGCGG	60	
DB	1	AAGTCTAAAGCCTTCGAGGCGCAAGGCGCTGCTACTGCGGCGGCTGCTTCTTAGTGCGG	60	
QY	61	GCTTCCGCGCCTTGCGTGTCTACGCGGCGCGCGCCGAGGAAGCCACTCGCAACCAAGGACCG	120	
DB	61	GCTTCCGCGCCTTGCGTGTCTACGCGGCGCGCGCCGAGGAAGCCACTCGCAACCAAGGACCG	120	
QY	121	GAGTGGAGGCGCGCAGCATGAGCGCGCAGCGCCGCTCCATAGCGACGCTCGGAGCGG	180	
DB	121	GAGTGGAGGCGCGCAGCATGAGCGCGCAGCGCCGCTCCATAGCGACGCTCGGAGCGG	180	
QY	181	TCGCGGCGGGGCGCGGGGGAAGGAAAATGCAACATGGCAGCAGCAATGAAACAGACAG	240	
DB	181	TCGCGGCGGGGCGCGGGGGAAGGAAAATGCAACATGGCAGCAGCAATGAAACAGACAG	240	
QY	241	CTGGGTGTTGAGATATTTGAAACTGCGGCACTGTGAGGAGAAATATTGAATTCACAGGATCGG	300	
DB	241	CTGGGTGTTGAGATATTTGAAACTGCGGCACTGTGAGGAGAAATATTGAATTCACAGGATCGG	300	
QY	301	CTTAAATTGGAGCCCTTTTATGTTGAGCGGTAATCTCTGGAGTCAGCTTAAAAAGCTGCTT	360	
DB	301	CTTAAATTGGAGCCCTTTTATGTTGAGCGGTAATCTCTGGAGTCAGCTTAAAAAGCTGCTT	360	
QY	361	GCCTGATACCAGAAAATATCATGGCTACATGATGGCTAAGGSCCACCATGATTTTCATGTTT	420	
DB	361	GCCTGATACCAGAAAATATCATGGCTACATGATGGCTAAGGSCCACCATGATTTTCATGTTT	420	
QY	421	GTGAAGAGGAATGATTCAGATGGACCTCATTTCCAGACAGAAATCTATTACCTTGGCATGTCT	480	
DB	421	GTGAAGAGGAATGATTCAGATGGACCTCATTTCCAGACAGAAATCTATTACCTTGGCATGTCT	480	
QY	481	GGTGAGAACAGAGAAAATACATCTGTTTATTTCTGAAAATTTCCCAAAAACCTATCATATGAGCA	540	
DB	481	GGTGAGAACAGAGAAAATACATCTGTTTATTTCTGAAAATTTCCCAAAAACCTATCAATAGACCA	540	

Qy	541	GCAGCTCTAAATGCTCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACCTGCGACTAT	600
Db	541	GCAGCTCTAAATGCTCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACCTGCGACTAT	600
Qy	601	GGATGTATTCTCGAGAGAGAACATTAAGAGGAAGAAACGATTTGGAAACAGTCGGA	660
Db	601	GGATGTATTCTCGAGAGAGAACATTAAGAGGAAGAAACGATTTGGAAACAGTCGGA	660
Qy	661	ATTGCTCTTACGATTATCACCAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGG	720
Db	661	ATTGCTCTTACGATTATCACCAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGG	720
Qy	721	ATTATACGTAAGAAAGTGGAGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAAT	780
Db	721	ATTATACGTAAGAAAGTGGAGGCCACAAGGATTTACGCAACAACTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTTGTCGCAACATACGAGTGGATTCGAAATTTATGCCCCGCTGATCCA	840
Db	781	CTAGTGGAAACTAGTTGTCGCAACATACGAGTGGATTCGAAATTTATGCCCCGCTGATCCA	840
Qy	841	GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTGAACAGAA	900
Db	841	GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTGAACAGAA	900
Qy	901	GAAAGAGACTCAGTTATGTCACATGAGCTAGCCAACTGGAGAGAGATGCCAGATCA	960
Db	901	GAAAGAGACTCAGTTATGTCACATGAGCTAGCCAACTGGAGAGAGATGCCAGATCA	960
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGAAAGAAATTTGATAGATATTTCTGGCTATTGTTGG	1020
Db	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGAAAGAAATTTGATAGATATTTCTGGCTATTGTTGG	1020
Qy	1021	TGTCAAAAGCTGAAACAACTCCAGCTGGTGGTGAATAATCTTAGAATTTCTATATGAAGAA	1080
Db	1021	TGTCAAAAGCTGAAACAACTCCAGCTGGTGGTGAATAATCTTAGAATTTCTATATGAAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGGAAATTTATCTATGTTTACATCCCTATGTTGGAAACAAGGAGG	1140
Db	1081	AATGATGAATCTGAGGTGGAAATTTATCTATGTTTACATCCCTATGTTGGAAACAAGGAGG	1140
Qy	1141	CGAGATTCATTTCCGTTATCCTAACAAGGTAACGGAATCCTTAAGTCACTTTTAAAGATG	1200
Db	1141	CGAGATTCATTTCCGTTATCCTAACAAGGTAACGGAATCCTTAAGTCACTTTTAAAGATG	1200
Qy	1201	TCAGAAATAATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATAAGGAACATAAT	1260
Db	1201	TCAGAAATAATGATTGATGCTGAAGGAAGATCATAGATGTCATAGATAAGGAACATAAT	1260
Qy	1261	CAACCTTTTGAGATTCTATTTTGAAGGAGTTGAATATATTGGCAGAGCTGGATGGACTCCT	1320
Db	1261	CAACCTTTTGAGATTCTATTTTGAAGGAGTTGAATATATTGGCAGAGCTGGATGGACTCCT	1320
Qy	1321	GAGGAAATAATGCTTGGTCCATCTCTACATAGATCGCTCCAGACTCGCTACAGATAGTG	1380
Db	1321	GAGGAAATAATGCTTGGTCCATCTCTACATAGATCGCTCCAGACTCGCTACAGATAGTG	1380
Qy	1381	TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGCAAGACTC	1440
Db	1381	TTGATCTCACTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGCAAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCTGATTTCTGTGCGGACCTAAATTTATCTATGAAGAAACACAGACATC	1500
Db	1441	ATTGAGTCAGTGCCTGATTTCTGTGCGGACCTAAATTTATCTATGAAGAAACACAGACATC	1500
Qy	1501	TGGAATAATTCATGACATCTTTTCATGTTTTTCCCAAGTCTACGAGAGGAAATTGAG	1560
Db	1501	TGGAATAATTCATGACATCTTTTCATGTTTTTCCCAAGTCTACGAGAGGAAATTGAG	1560
Qy	1561	TTTATTTTGGCTCTGAAATGCAAAAACAGGTTTTCCGTCATTTTATACAAAATTCATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAAATGCAAAAACAGGTTTTCCGTCATTTTATACAAAATTCATCTATT	1620
Qy	1621	TTAAGGAAGCAAAATAAAACGATCCAGTGGTGGGCTGCTGCTCCAAGTGAATTTCAAG	1680

```
Db 1621 TTAAGGAAAGCAAAATATAAAGATCCAGTGTGGGCTGCCTGCTCCAGTGATTTTCAAG 1680
Qy TGTCTTATCAAGAGGAGATAGCAATATACAGTGTGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
Db TGTCTTATCAAGAGGAGATAGCAATATACAGTGTGTGAATGGGAAGTTCTTTGGCCGGCAT 1740
Qy GGATCTAATATCAAGAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db GGATCTAATATCAAGAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Qy TCCCTTTTAGAGATCACCTGTAGTACGTAGTACGTAAATCCTGGAGAGGTCACAAGG 1860
Db TCCCTTTTAGAGATCACCTGTAGTACGTAGTACGTAAATCCTGGAGAGGTCACAAGG 1860
Qy CTGACTCACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATA 1920
Db CTGACTCACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATA 1920
Qy AGTAAGTATAGTAACAGAGATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCCT 1980
Db AGTAAGTATAGTAACAGAGATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCCT 1980
Qy GAAGATCAACCACTTGCACCAAGAAATTTTGGGCCACCATTTTGGATTCAGCAGGT 2040
Db GAAGATCAACCACTTGCACCAAGAAATTTTGGGCCACCATTTTGGATTCAGCAGGT 2040
Qy CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACCA 2100
Db CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTACCA 2100
Qy TTGTATGGGATGCTTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
Db TTGTATGGGATGCTTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG 2160
Qy CTGTTTCATATATGGTGGTC 2179
Db CTGTTTCATATATGGTGGTC 2179

RESULT 7
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junier, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-22

Query Match 63.6%; Score 1984; DB 3; Length 4685;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2034; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAGTGTCTAAAGCCTCCGAGGCCAAGGCCGCTCTACTACTCGCCGCGCTCTTCTTAGTGCGG 60
Db 1 AAGTGTCTAAAGCCTCCGAGGCCAAGGCCGCTCTACTACTCGCCGCGCTCTTCTTAGTGCGG 60
Qy 61 CGTTCCGCGCTGGGTGTGTCACGCGCGCGCGCGCGAGGAAAGCACTGCAACACGAGACCG 120
```

```
Db 61 CGTTCCGCGCTGGGTGTGTCACGCGCGCGCGCGCGAGGAAAGCACTGCAACACGAGACCG 120
Qy 121 GAGTGGAGGCGCGCAGCATGAAGCGGCGCAGGCCCGCTCCATAGCGCACGTCCGGACGG 180
Db 121 GAGTGGAGGCGCGCAGCATGAAGCGGCGCAGGCCCGCTCCATAGCGCACGTCCGGACGG 180
Qy 181 TCCGGGCGGCGCGCGGGAAGAAATGCAACATGCGCAGCAGCAATGGAACAGAAACAG 240
Db 181 TCCGGGCGGCGCGCGGGAAGAAATGCAACATGCGCAGCAGCAATGGAACAGAAACAG 240
Qy 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGG 300
Qy 301 CTTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGGTACAGTTTAAAGAGCTGCTT 360
Db 301 CTTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGGTACAGTTTAAAGAGCTGCTT 360
Qy 361 GCCGATACAGAAATATCATGCTACATGATGGCTTAAGGCCACACATGATTTTCATGTTT 420
Db 361 GCCGATACAGAAATATCATGCTACATGATGGCTTAAGGCCACACATGATTTTCATGTTT 420
Qy 421 GTGAAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAACTATACCTTGGCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAACTATACCTTGGCATGCT 480
Qy 481 GGTGAGAAACAGAGAAATACACTGTTTATTCGAAATTCGAAATTCGAAATTCATATAGCA 540
Db 481 GGTGAGAAACAGAGAAATACACTGTTTATTCGAAATTCGAAATTCGAAATTCATATAGCA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
Qy 601 GGAATGTATTTCTCGAGAAGAAACTATTAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
Db 601 GGAATGTATTTCTCGAGAAGAAACTATTAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
Qy 661 ATTGCTTCTTACGATTTACCAAGAGAGTGGAAATTTCTGTTTCAAGCGCGGTAGTGA 720
Db 661 ATTGCTTCTTACGATTTACCAAGAGAGTGGAAATTTCTGTTTCAAGCGCGGTAGTGA 720
Qy 721 ATTTATCACGTAAAGATGGAGGCCACAAGGATTTAGCAACAACTTTTAAGGCCCAAT 780
Db 721 ATTTATCACGTAAAGATGGAGGCCACAAGGATTTAGCAACAACTTTTAAGGCCCAAT 780
Qy 781 CTAGTGGAACTAGTTGTCCACACATACGATGGATCCAAATTTATGCCCTGCTGATCCA 840
Db 781 CTAGTGGAACTAGTTGTCCACACATACGATGGATCCAAATTTATGCCCTGCTGATCCA 840
Qy 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACAGAGAA 900
Db 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACAGAGAA 900
Qy 901 GAAAGGAGACTCATTTATGTGCNCAATGAGCTAGGCCAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCATTTATGTGCNCAATGAGCTAGGCCAACATGGAAGAGATGCCAGATCA 960
Qy 961 GCTGGAGTCTGCTACCTTTGTTCTTCAAGAGAAATTTGATAGATATCTTGGCTATTTGGTGG 1020
Db 961 GCTGGAGTCTGCTACCTTTGTTCTTCAAGAGAAATTTGATAGATATCTTGGCTATTTGGTGG 1020
Qy 1021 TGTCCAAAGCTGAAAACAACTCCAGTGTGTGTAATAATTTCTTGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAGCTGAAAACAACTCCAGTGTGTGTAATAATTTCTTGAATTTCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGGAATAATTTATCATGTTTACATCCCTTATGTTGGAACAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAATAATTTATCATGTTTACATCCCTTATGTTGGAACAGGAGG 1140
Qy 1141 GCAGATTTCATTCGTTTATCTTAAACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTTCATTCGTTTATCTTAAACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAAGATG 1200
```



[illegible]

## RESULT 8

```

US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. 6844160
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PRO
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976-674-14
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240
; PRIOR FILING DATE: 2000-10-12

```

Db 2130 TACATGGTTCCTGGATGAGAAATGTCATTTTGGACATACCGATATATTATCTGAGTTTT 2189  
Qy 2708 TAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCAGGAGACACAGCATAA 2767  
Db 2190 TAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCCAGGAGACACAGCATAA 2249  
Qy 2768 GAGTTCCTGAAATCGGAGAAACATATGAACTGCACTCTTTTGGACATACCTTCAGAAACC 2827  
Db 2250 GAGTTCCTGAAATCGGAGAAACATATGAACTGCACTCTTTTGGACATACCTTCAGAAACC 2309  
Qy 2828 TTGATCACGTAATGCTGCTCTAAAGAGTATATAATTTTGAACCTGTGTGAGAACTCTCTGG 2887  
Db 2310 TTGATCACGTAATGCTGCTCTAAAGAGTATATAATTTTGAACCTGTGTGAGAACTCTCTGG 2369  
Qy 2888 TATACACTGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTGATC 2947  
Db 2370 TATACACTGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTGATC 2429  
Qy 2948 ATACATATTTGATACCTGCCATGTAACTACTCTGAAATAAATGTTGGTGCATGCA 3007  
Db 2430 ATACATATTTGATACCTGCCATGTAACTACTCTGAAATAAATGTTGGTGCATGCA 2489  
Qy 3008 GGGGTCTACGGTTTGGTGTAGTAACTAACTACTCTAAACCCACATGCTCAAAATCAAAATG 3067  
Db 2490 GGGGTCTACGGTTTGGTGTAGTAACTAACTACTCTAAACCCACATGCTCAAAATCAAAATG 2549  
Qy 3068 ATACATATTTCTGAGAGACCCAGCAATACCAATGAATTTACTTAAACCAAAAAA 3120  
Db 2550 ATACATATTTCTGAGAGACCCAGCAATACCAATGAATTTACTTAAACCAAAAAA 2602

RESULT 9  
US-10-070-464-6  
; Sequence 6, Application US/10070464  
; Patent No. 6841564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT;  
; APPLICANT: GORRELL, Mark Douglas  
; FILE OF INVENTION: DIPETIDYL PEPTIDASES  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1669  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-070-464-6

Query Match 32.8%; Score 1023; DB 3; Length 1669;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1023; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1164 AACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGA 1223  
Db 1 AACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGA 60  
Qy 1224 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAACTCAACCTTTTGAAGATTTCTATTGA 1283  
Db 61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAACTCAACCTTTTGAAGATTTCTATTGA 120  
Qy 1284 AGGAGTGTATATTTGACAGAGCTGGAGTGGTCTGAGGAAATATGCTGGTCCAT 1343  
Db 121 AGGAGTGTATATTTGACAGAGCTGGAGTGGTCTGAGGAAATATGCTGGTCCAT 180

Qy 1344 CTTACTAGATGCTCCAGACTCGCTACAGATAGTGTGATCTCACCCTGAATTTAT 1403  
Db 181 CTTACTAGATGCTCCAGACTCGCTACAGATAGTGTGATCTCACCCTGAATTTAT 240  
Qy 1404 CCCAGTGAAGATGATGTTATGGAAAGGACAGACTCATTTAGTCAAGTGCCTGATCTGT 1463  
Db 241 CCCAGTGAAGATGATGTTATGGAAAGGACAGACTCATTTAGTCAAGTGCCTGATCTGT 300  
Qy 1464 GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATPAATAATCTCCATGACATCTT 1523  
Db 301 GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATPAATAATCTCCATGACATCTT 360  
Qy 1524 TCATGTTTTTCCCAAGATCACGAAGAGAAATTTGAGTTTATTTTGGCTCTGAATGCAA 1583  
Db 361 TCATGTTTTTCCCAAGATCACGAAGAGAAATTTGAGTTTATTTTGGCTCTGAATGCAA 420  
Qy 1584 AACAGGTTTCCCGTCAATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAACG 1643  
Db 421 AACAGGTTTCCCGTCAATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAACG 480  
Qy 1644 ATCCAGTGGTGGCTGCTGCTCCAAAGTCAATTTCAAGTGTCTCTATCAAAAGAGGAGATAGC 1703  
Db 481 ATCCAGTGGTGGCTGCTGCTCCAAAGTCAATTTCAAGTGTCTCTATCAAAAGAGGAGATAGC 540  
Qy 1704 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGACATGATCTTAATATCTCAAGTTGATGA 1763  
Db 541 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGACATGATCTTAATATCTCAAGTTGATGA 600  
Qy 1764 AGTCAGAGGCTGATATATTTTGAAGGACCAAGACTCCCTTTAGAGCATACCTGTGA 1823  
Db 601 AGTCAGAGGCTGATATATTTTGAAGGACCAAGACTCCCTTTAGAGCATACCTGTGA 660  
Qy 1824 CGTAGTCAGTACGTAAATCCCTGGAGAGGTGACAGGCTGACCTGACCTGCTACTCACA 1883  
Db 661 CGTAGTCAGTACGTAAATCCCTGGAGAGGTGACAGGCTGACCTGACCTGCTACTCACA 720  
Qy 1884 TTCTTCTGATCAGTCAAGCATGCTGCACTTTTATATAGTATAGTATAGTACCAAGAA 1943  
Db 721 TTCTTCTGATCAGTCAAGCATGCTGCACTTTTATATAGTATAGTATAGTACCAAGAA 780  
Qy 1944 TCCACACTGTGCTCCTTTTCAAGTATCAAGTCTCAAGATGATGCCAACTTGCAGAAC 2003  
Db 781 TCCACACTGTGCTCCTTTTCAAGTATCAAGTCTCAAGATGATGCCAACTTGCAGAAC 840  
Qy 2004 AAAGAAATTTTGGGCCACCACTTTTGGATTCAGCAGGCTCTCTCTGCTACTATCTCTCC 2063  
Db 841 AAAGAAATTTTGGGCCACCACTTTTGGATTCAGCAGGCTCTCTCTGCTACTATCTCTCC 900  
Qy 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGAATGCTCTCAAGCC 2123  
Db 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGAATGCTCTCAAGCC 960  
Qy 2124 TCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATGTTGTTCTCTCA 2183  
Db 961 TCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATGTTGTTCTCTCA 1020  
Qy 2184 GGT 2186  
Db 1021 GGT 1023

RESULT 10  
US-09-976-674-10  
; Sequence 10, Application US/09976674  
; Patent No. 6841180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinesya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674

; CURRENT FILING DATE: 2001-10-12									
; PRIOR APPLICATION NUMBER: US 60/240,117									
; PRIOR FILING DATE: 2000-10-12									
; NUMBER OF SEQ ID NOS: 61									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 10									
; LENGTH: 1356									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-976-674-10									
Query Match 28.1%; Score 877; DB 3; Length 1356;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 927; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	AAGTGTAAAGCCTCGAGGCGCAAGCGCGCTGTCTACTGCGCGCGCTGCTTCTTAGTGCGG	60						
DB	1	AAGTGTAAAGCCTCGAGGCGCAAGCGCGCTGTCTACTGCGCGCGCTGCTTCTTAGTGCGG	60						
QY	61	CGTTGCGCCCTGGGTGTACCCGCGCGCGCGGAGGAGCCACTGCAACCGAGCGG	120						
DB	61	CGTTGCGCCCTGGGTGTACCCGCGCGCGCGGAGGAGCCACTGCAACCGAGCGG	120						
QY	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGAGCG	180						
DB	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGAGCG	180						
QY	181	TCCGGCGGGCGCGGGGGAAGAAAATGCAACATGGCAGGAGCAATGGAACAGAACAG	240						
DB	181	TCCGGCGGGCGCGGGGGAAGAAAATGCAACATGGCAGGAGCAATGGAACAGAACAG	240						
QY	241	CTGGGTGTGAGATATTTGAACCTGCGGACTGTGAGGAGATATTTGAATCACAAGGATCG	300						
DB	241	CTGGGTGTGAGATATTTGAACCTGCGGACTGTGAGGAGATATTTGAATCACAAGGATCG	300						
QY	301	CTAAATTTGAGCGCTTTTATCTGTAGCGGTATTTCTGGAGTCAAGCTTAAAAAGCTGCTT	360						
DB	301	CTAAATTTGAGCGCTTTTATCTGTAGCGGTATTTCTGGAGTCAAGCTTAAAAAGCTGCTT	360						
QY	361	GCCGATACCAAGAAAATATCATGGCTACATGATGGCTAAGGACACATGATTTTCATGTTT	420						
DB	361	GCCGATACCAAGAAAATATCATGGCTACATGATGGCTAAGGACACATGATTTTCATGTTT	420						
QY	421	GTCAAGAGGATGATCCAGATGGACTCATTCAGACAGAACTCTATTACTTGGCATGTCT	480						
DB	421	GTCAAGAGGATGATCCAGATGGACTCATTCAGACAGAACTCTATTACTTGGCATGTCT	480						
QY	481	GGTGAGAACAGAAAAATACACTGTTTTTATCTGAAATTTCCAAAACTATCAATAGAGCA	540						
DB	481	GGTGAGAACAGAAAAATACACTGTTTTTATCTGAAATTTCCAAAACTATCAATAGAGCA	540						
QY	541	GCAATCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600						
DB	541	GCAATCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600						
QY	601	GGAATGTAATCTCGAGAGGAACTATTAAGAGAAAGAAAACGATTTGGAACAGTCGGA	660						
DB	601	GGAATGTAATCTCGAGAGGAACTATTAAGAGAAAGAAAACGATTTGGAACAGTCGGA	660						
QY	661	ATTGCTTCTTACGATATCACCAAGGAGTGGAACTTTCTGTTTCAAGCGGTAGTGGGA	720						
DB	661	ATTGCTTCTTACGATATCACCAAGGAGTGGAACTTTCTGTTTCAAGCGGTAGTGGGA	720						
QY	721	ATTATATCACTAAAGATGAGGGCCACAAAGATTTAGCAACACCTTTTAAGGCCCAAT	780						
DB	721	ATTATATCACTAAAGATGAGGGCCACAAAGATTTAGCAACACCTTTTAAGGCCCAAT	780						
QY	781	CTAGTGGAACTAGTGTCTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840						
DB	781	CTAGTGGAACTAGTGTCTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840						
QY	841	GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAACAGAGAA	900						
DB	841	GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAACAGAGAA	900						
; Sequence 4, Application US/10070464									
; Patent No. 6881564									
; GENERAL INFORMATION:									
; APPLICANT: ABBOTT, Catherine Anne									
; APPLICANT: CORRELL, Mark Douglas									
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES									
; FILE REFERENCE: GH-007									
; CURRENT APPLICATION NUMBER: US/10/070,464									
; CURRENT FILING DATE: 2002-03-07									
; PRIOR APPLICATION NUMBER: PCT/AU00/01085									
; PRIOR FILING DATE: 2000-09-11									
; PRIOR APPLICATION NUMBER: AU PQ5709									
; PRIOR FILING DATE: 2000-02-18									
; PRIOR APPLICATION NUMBER: AU PQ2762									
; PRIOR FILING DATE: 1999-09-10									
; NUMBER OF SEQ ID NOS: 8									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 4									
; LENGTH: 1197									
; TYPE: DNA									
; ORGANISM: Homo Sapiens									
US-10-070-464-4									
Query Match 25.3%; Score 790; DB 3; Length 1197;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	2331	GGGTCAAAATAGAAATTCAGGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA	2390						
DB	404	GGGTCAAAATAGAAATTCAGGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA	463						
QY	2391	TTTCATTGACTTAGATCGGTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCT	2450						
DB	464	TTTCATTGACTTAGATCGGTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCT	523						
QY	2451	GATGGCAATTAATGCAGAGGTCCAGATATCTTCAAGGTGCTATTTGCTGGGGCCCCAGTCC	2510						
DB	524	GATGGCAATTAATGCAGAGGTCCAGATATCTTCAAGGTGCTATTTGCTGGGGCCCCAGTCC	583						
QY	2511	TCTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATGGGTCACTGACCCAGAA	2570						
DB	584	TCTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATGGGTCACTGACCCAGAA	643						
QY	2571	TGAACAGGGCTATTTACTTAGGATCTGTGACCATGCAAGCAGAAAAAGTTCCCTCTGAACC	2630						
DB	644	TGAACAGGGCTATTTACTTAGGATCTGTGACCATGCAAGCAGAAAAAGTTCCCTCTGAACC	703						
QY	2631	AAATCGTTTACTGCTCTTACATGTTTTCTGGATGAGAATGTCATTTTGGACATACCAG	2690						
DB	704	AAATCGTTTACTGCTCTTACATGTTTTCTGGATGAGAATGTCATTTTGGACATACCAG	763						
QY	2691	TATATTACTGAGTTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA	2750						
DB	764	TATATTACTGAGTTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA	823						
QY	2751	GGAGACACAGCATTAAGAGTTCTCTGAATCGGGAGAACATTTATGAATGCACTTTTGGCA	2810						
DB	824	GGAGACACAGCATTAAGAGTTCTCTGAATCGGGAGAACATTTATGAATGCACTTTTGGCA	883						
QY	2811	CTACCTTCAAGAAAACTTGGATCAAGTATTTGCTGCTCTAAAGTGATATAATTTTGACC	2870						
DB	884	CTACCTTCAAGAAAACTTGGATCAAGTATTTGCTGCTCTAAAGTGATATAATTTTGACC	943						
QY	2871	TGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAACTCAACAG	2930						

Db 944 TGTGTAGACTCTCTGTGTATACACTGGCTATTATTAACCAATGAGAGGTTTAAATCAACAG 1003  
Qy 2931 AAAACACAGAAATGATCATACATTTTGTATACCTGCCATGTAAACATCTACTCTCTGAAAT 2990  
Db 1004 AAAACACAGAAATGATCATACATTTTGTATACCTGCCATGTAAACATCTACTCTCTGAAAT 1063  
Qy 2991 AAATGTGGTCCATGAGGGGCTACGGTTTGTGTAGTAATCTAAATACCTTTAACCCAC 3050  
Db 1064 AAATGTGGTCCATGAGGGGCTACGGTTTGTGTAGTAATCTAAATACCTTTAACCCAC 1123  
Qy 3051 ATGCTCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATACCAATGAATTA 3110  
Db 1124 ATGCTCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATACCAATGAATTA 1183  
Qy 3111 AAAAAAAAAA 3120  
Db 1184 AAAAAAAAAA 1193

RESULT 12  
US-10-070-464-8  
; Sequence 8, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: CORRELL, Mark Douglas  
; TITLE OF INVENTION: DIFEPIIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 1083  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-070-464-8

Query Match 25.3%; Score 789; DB 3; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1755 AGTTGATGAGTCAGAGGCTGTATATTTGAAGGACCAAGAGACTCCCTTTAGAGCA 1814  
Db 295 AGTTGATGAGTCAGAGGCTGTATATTTGAAGGACCAAGAGACTCCCTTTAGAGCA 354  
Qy 1815 TCACCTGTACGTAGTCAGTACGTAAATCTCTGAGAGGTGCAAGGCTGACTGACCGTGG 1874  
Db 355 TCACCTGTACGTAGTCAGTACGTAAATCTCTGAGAGGTGCAAGGCTGACTGACCGTGG 414  
Qy 1875 TACTACATTTCTGTGTCATAGTCAGCACTGTGACTTTTATAAGTAAGTATAGTAA 1934  
Db 415 TACTACATTTCTGTGTCATAGTCAGCACTGTGACTTTTATAAGTAAGTATAGTAA 474  
Qy 1935 CAGAAGAAATCCACATGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAAC 1994  
Db 475 CAGAAGAAATCCACATGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAAC 534  
Qy 1995 TTGCAAAACAAAGAAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTTGACTA 2054  
Db 535 TTGCAAAACAAAGAAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTTGACTA 594  
Qy 2055 TACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGATTTACATTTGATGGATGCT 2114  
Db 595 TACTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGATTTACATTTGATGGATGCT 654

Qy 2115 CTCAAGCCTCATGATCTACAGCCTCGAAAGAAATATCTCTACTGTCTGTTCATATATGG 2174  
Db 655 CTCAAGCCTCATGATCTACAGCCTCGAAAGAAATATCTCTACTGTCTGTTCATATATGG 714  
Qy 2175 TGGTCCCTCAGGTGAGTTCGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGA 2234  
Db 715 TGGTCCCTCAGGTGAGTTCGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGA 774  
Qy 2235 TACCTAGCCTCTTAGTGTATGTGTGTAGTATAGTATAGCAACAGGGGATCTCTGTCACCG 2294  
Db 775 TACCTAGCCTCTCTAGTGTATGTGTGTAGTATAGCAACAGGGGATCTCTGTCACCG 834  
Qy 2295 AGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATAGAAATGACGATCA 2354  
Db 835 AGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATAGAAATGACGATCA 894  
Qy 2355 GGTGGAAGGACTCCAATATCTAGCTTCTCGATATATTTTCAATGACTTAGATCGTGGG 2414  
Db 895 GGTGGAAGGACTCCAATATCTAGCTTCTCGATATATTTTCAATGACTTAGATCGTGGG 954  
Qy 2415 CATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGA 2474  
Db 955 CATCCACGGCTGCTCTATGAGGATACCTCTCTCCCTGATGGCATTAAATGCAGAGGTGAGA 1014  
Qy 2475 TATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAG 2534  
Db 1015 TATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAG 1074  
Qy 2535 ATACACGGA 2543  
Db 1075 ATACACGGA 1083

RESULT 13  
US-09-976-674-18  
; Sequence 18, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Junien, Jean-Louis  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 832  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-18

Query Match 24.4%; Score 760; DB 3; Length 832;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGTCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTCTTAGTCCG 60  
Db 1 AAGTCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTCTTAGTCCG 60  
Qy 61 GGTTCGCGCGCTGGTTGTCTACCGCGCCGCCCGCCAGGAGCACTGCAACACAGACCG 120  
Db 61 GGTTCGCGCGCTGGTTGTCTACCGCGCCGCCCGCCAGGAGCACTGCAACACAGACCG 120  
Qy 121 GAGTGGAGCGCGCCAGCATGAAGAGCGCGCAGCGCCGCTCCATAGCGCACGTCGGGACG 180  
Db 121 GAGTGGAGCGCGCGCAGCATGAAGAGCGCGCAGCGCCGCTCCATAGCGCACGTCGGGACG 180  
Qy 181 TCCGGCGCGCGCGCGGGGGAAGGAAATGCAACATGGCAGCAATGGAACAGACAG 240



Mon May 8 13:41:52 2006

us-10-825-632-2.oligo.rni

Page 16

DB 335 AGAATGTCATTGTCACATACCAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAGC 394

Search completed: May 3, 2006, 02:28:13  
Job time : 940 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 09:10:06 ; Search time 2192 Seconds  
(without alignments)  
11770.283 Million cell updates/sec

Title: US-10-825-632-2  
Perfect score: 3120  
Sequence: 1 aagtgtaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 413469005 residues

Total number of hits satisfying chosen parameters: 19584538

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA\_Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3120	100.0	3120	7	US-10-415-122-5
2	3120	100.0	3120	8	US-10-825-632-2
3	3003	96.2	3106	7	US-10-311-035-30
4	2957	94.8	3143	6	US-10-170-789-37
5	2889	92.6	4829	3	US-09-976-674-12
6	2889	92.6	4829	3	US-10-982-512-12
7	2547	81.6	2649	6	US-10-170-789-39
8	2496	80.0	2649	6	US-10-054-776-1
9	2495	80.0	2671	3	US-09-976-674-2
10	2495	80.0	2671	9	US-10-982-512-2
11	2135	68.4	2830	9	US-10-956-157-2177
12	2128	68.2	4523	3	US-09-976-674-8
13	2128	68.2	4523	9	US-10-982-512-8
14	2128	68.2	4676	3	US-09-976-674-20
15	2128	68.2	4676	3	US-10-982-512-20
16	1984	63.6	4685	3	US-09-976-674-22
17	1984	63.6	4685	9	US-10-982-512-22
18	1191	38.2	4309	3	US-09-976-674-14
19	1191	38.2	4309	9	US-10-982-512-14
20	1164	37.3	2510	7	US-10-275-505-16
21	1164	37.3	2510	10	US-11-140-224-16
22	1023	32.8	1669	8	US-10-825-632-6
23	877	28.1	1356	3	US-09-976-674-10

24	877	28.1	1356	9	US-10-982-512-10	Sequence 10, Appl
25	790	25.3	1197	8	US-10-825-632-4	Sequence 4, Appl
26	789	25.3	1083	8	US-10-825-632-8	Sequence 8, Appl
27	760	24.4	832	3	US-09-976-674-18	Sequence 18, Appl
28	760	24.4	832	3	US-10-982-512-18	Sequence 18, Appl
29	737	23.6	925	6	US-10-264-237-710	Sequence 710, Appl
30	590	18.9	600	9	US-10-956-157-7412	Sequence 7412, Ap
31	472	15.1	620	3	US-09-976-674-16	Sequence 16, Appl
32	472	15.1	620	9	US-10-982-512-16	Sequence 16, Appl
33	431	13.8	502	3	US-09-918-995-19585	Sequence 19585, A
34	254	8.1	561	3	US-09-764-891-877	Sequence 877, App
35	167	5.4	168	9	US-10-467-851-254	Sequence 254, App
36	156	5.0	4797	3	US-09-764-891-7074	Sequence 7074, Ap
37	118	3.8	1346	4	US-09-925-065A-669313	Sequence 669313,
38	112	3.6	609	4	US-09-925-065A-743558	Sequence 743558,
39	112	3.6	609	4	US-09-925-065A-743559	Sequence 743559,
40	73	2.3	612	4	US-09-925-065A-818446	Sequence 818446,
41	46	1.5	60	3	US-09-908-975-15085	Sequence 15085, A
42	30	1.0	30	8	US-10-825-632-27	Sequence 27, Appl
43	28	0.9	497	4	US-09-925-065A-482467	Sequence 482467,
44	27	0.9	392	3	US-09-783-590-879	Sequence 879, App
45	25	0.8	25	9	US-10-956-157-58438	Sequence 58438, A

ALIGNMENTS

RESULT 1  
US-10-415-122-5  
; Sequence 5, Application US/10415122  
; Publication No. US2004003369A1  
; GENERAL INFORMATION:  
; APPLICANT: THE UNIVERSITY OF SYDNEY  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: FPI5217  
; CURRENT APPLICATION NUMBER: US/10/415,122  
; CURRENT FILING DATE: 2003-08-07  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 5  
; LENGTH: 3120  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-415-122-5

Query Match		100.0%;	Score 3120;	DB 7;	Length 3120;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 3120;		Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Qy	1	AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGC	60		
Db	1	AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGC	60		
Qy	61	CGTTCCGCGCTGGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGCACACGAGCG	120		
Db	61	CGTTCCGCGCTGGGTGTGTCACCGCGCGCGCGAGGAGCCACTGCACACGAGCG	120		
Qy	121	GAGTGGAGCGCGGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGGA	180		
Db	121	GAGTGGAGCGCGGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGGA	180		
Qy	181	TCCGGCGCGCGCGGGGGAAGAAAATGCAATGGCAGCAGCAATGGAACAGAAC	240		
Db	181	TCCGGCGCGCGCGGGGGAAGAAAATGCAATGGCAGCAGCAATGGAACAGAAC	240		
Qy	241	CTGGGTGTGAGATTTTGAATCTGCGGACTGTGAGGAGATTTGAATACAGGATCG	300		
Db	241	CTGGGTGTGAGATTTTGAATCTGCGGACTGTGAGGAGATTTGAATACAGGATCG	300		
Qy	301	CCTAAATTGGAGCCTTTTATGTTAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGT	360		
Db	301	CCTAAATTGGAGCCTTTTATGTTAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGT	360		

QY 361 GCCGATACCAGAAAATATCATGGCTACATGATGGCTAAGGCACACACATGATTTTCATGTTT 420  
Db 361 GCCGATACCAGAAAATATCATGGCTACATGATGGCTAAGGCACACACATGATTTTCATGTTT 420  
QY 421 GTGAAGAGGATGATCCAGATGAGACTCATTCAGACAGATCTATTACCTTGGCCATGCT 480  
Db 421 GTGAAGAGGATGATCCAGATGAGACTCATTCAGACAGATCTATTACCTTGGCCATGCT 480  
QY 481 GGTGAGACACAGAGAAAATACACTGTTTTTATCTCGAAATTTCCCAAACTATCAATAGAGCA 540  
Db 481 GGTGAGACACAGAGAAAATACACTGTTTTTATCTCGAAATTTCCCAAACTATCAATAGAGCA 540  
QY 541 GCAGCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
Db 541 GCAGCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
QY 601 GGAATGTATCTCGAGAGAGAAACTATTAAAGAGAGAGAAACGCAATTTGGAACAGTCCGA 660  
Db 601 GGAATGTATCTCGAGAGAGAAACTATTAAAGAGAGAGAAACGCAATTTGGAACAGTCCGA 660  
QY 661 ATTGCTTTCTTACGATTTATCAACGAAGAGTGAACATTTCTGTTTTCAAGCCGGTAGTGA 720  
Db 661 ATTGCTTTCTTACGATTTATCAACGAAGAGTGAACATTTCTGTTTTCAAGCCGGTAGTGA 720  
QY 721 ATTTATCAGCTAAAAGATGAGGGCCCAAGGATTTTACGCAACACCTTTTAAAGCCCAAT 780  
Db 721 ATTTATCAGCTAAAAGATGAGGGCCCAAGGATTTTACGCAACACCTTTTAAAGCCCAAT 780  
QY 781 CTAGTGAACACTAGTTGTCTCCACATACCGATGGATCCCAAAATTTATGCCCGCTGATCCA 840  
Db 781 CTAGTGAACACTAGTTGTCTCCACATACCGATGGATCCCAAAATTTATGCCCGCTGATCCA 840  
QY 841 GACTGGATTCCTTTTATACATAGCAACGATTTTGGATCTTACATCTGTAACAGAGAA 900  
Db 841 GACTGGATTCCTTTTATACATAGCAACGATTTTGGATCTTACATCTGTAACAGAGAA 900  
QY 901 GAAAGGAGACTCATTTATGTGCACAAATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960  
Db 901 GAAAGGAGACTCATTTATGTGCACAAATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960  
QY 961 GCTGGAGTCTGCTACCTTTGTTCTTCAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020  
Db 961 GCTGGAGTCTGCTACCTTTGTTCTTCAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020  
QY 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTGATAAAATTTCTTAGAATTTCTATAGAAGAA 1080  
Db 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGTGATAAAATTTCTTAGAATTTCTATAGAAGAA 1080  
QY 1081 AATGATGAATCTGAGGTGAAATTTATTCATGTTACATCCCTATGTTGGAAACAAAGGAGG 1140  
Db 1081 AATGATGAATCTGAGGTGAAATTTATTCATGTTACATCCCTATGTTGGAAACAAAGGAGG 1140  
QY 1141 GCAGATTCATTCGGTTATCCTTAAACAGGTACAGCAAAATCCTTAAAGTCACCTTTTAAAGATG 1200  
Db 1141 GCAGATTCATTCGGTTATCCTTAAACAGGTACAGCAAAATCCTTAAAGTCACCTTTTAAAGATG 1200  
QY 1201 TCAGAAATATGATTCATCTGAAGGAGGATCATAGATGTCATAGATAGGAACTAATTT 1260  
Db 1201 TCAGAAATATGATTCATCTGAAGGAGGATCATAGATGTCATAGATAGGAACTAATTT 1260  
QY 1261 CAACCTTTTGAATTTCTAATTTCAAGAGTTTGAATATATTGCGAGCTGATGGACTCT 1320  
Db 1261 CAACCTTTTGAATTTCTAATTTCAAGAGTTTGAATATATTGCGAGCTGATGGACTCT 1320  
QY 1321 GAGGAAATATGCTTTGGTCCATCTACTAGATCGCTCCAGACTCCGCTTACAGATGAGTGG 1380  
Db 1321 GAGGAAATATGCTTTGGTCCATCTACTAGATCGCTCCAGACTCCGCTTACAGATGAGTGG 1380  
QY 1381 TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC 1440  
Db 1381 TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGACTC 1440  
QY 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500

Db 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500  
QY 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAAGTCACGAGAGAGAAATTTGAG 1560  
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAAAGTCACGAGAGAGAAATTTGAG 1560  
QY 1561 TTTATTTTTCCTCTCGAATGCAAAACAGGTTTTCCGTCATTTTATACAAAATTTACATCTATT 1620  
Db 1561 TTTATTTTTCCTCTCGAATGCAAAACAGGTTTTCCGTCATTTTATACAAAATTTACATCTATT 1620  
QY 1621 TTAAGGAAAGCAAAATATAAAGATCGATCGGTGGGCTGCTGCTCCAAAGTGAATTTCAAG 1680  
Db 1621 TTAAGGAAAGCAAAATATAAAGATCGATCGGTGGGCTGCTGCTCCAAAGTGAATTTCAAG 1680  
QY 1681 TGTCTTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT 1740  
Db 1681 TGTCTTATCAAGAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTTGGCCGGCAT 1740  
QY 1741 GGATCTAATATCCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800  
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800  
QY 1801 TCCCCTTTTAGACATCACCTGTACGTAGTCAGTTAGCTTAAGTAAATCCTGGAGAGTGCACAGG 1860  
Db 1801 TCCCCTTTTAGACATCACCTGTACGTAGTCAGTTAGCTTAAGTAAATCCTGGAGAGTGCACAGG 1860  
QY 1861 CTGACTGACCGCTGGCTACTCAATTTCTGCTGCATCAGTCAGCATCTGTACTTCTTTATA 1920  
Db 1861 CTGACTGACCGCTGGCTACTCAATTTCTGCTGCATCAGTCAGCATCTGTACTTCTTTATA 1920  
QY 1921 AGTAGTATAGTAAACAGAGAAATCCACATCTGTGTCTCTTTTCAAGCTATCAAGTCTCT 1980  
Db 1921 AGTAGTATAGTAAACAGAGAAATCCACATCTGTGTCTCTTTTCAAGCTATCAAGTCTCT 1980  
QY 1981 GAAAGATGACCCAACTTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT 2040  
Db 1981 GAAAGATGACCCAACTTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT 2040  
QY 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCTGGAATTTTACA 2100  
Db 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCTGGAATTTTACA 2100  
QY 2101 TTGTATGGGATGCTCTTACAGGCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTG 2160  
Db 2101 TTGTATGGGATGCTCTTACAGGCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTG 2160  
QY 2161 CTGTTTATATATGTTGGTCTCTCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220  
Db 2161 CTGTTTATATATGTTGGTCTCTCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220  
QY 2221 TATTTCCGCTTGAATACCTACCTCTCTAGGTTATGTTGGTGTAGTAGAGACACAGG 2280  
Db 2221 TATTTCCGCTTGAATACCTACCTCTCTAGGTTATGTTGGTGTAGTAGAGACACAGG 2280  
QY 2281 GGATCTGTACCGAGGCTTAAATTTGAAGGCGCTTAAATATAAATGCGTCAATA 2340  
Db 2281 GGATCTGTACCGAGGCTTAAATTTGAAGGCGCTTAAATATAAATGCGTCAATA 2340  
QY 2341 GAAATTTGACGATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTGAC 2400  
Db 2341 GAAATTTGACGATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTGAC 2400  
QY 2401 TTAGATCTGTGGGCATCCAAGCTGTGCTTATGGAGGATACCTCTCCCTGATGGCAATTA 2460  
Db 2401 TTAGATCTGTGGGCATCCAAGCTGTGCTTATGGAGGATACCTCTCCCTGATGGCAATTA 2460  
QY 2461 ATCCAGAGGTTCAGATATCTTCAAGGTTTCTTATTTCTGGGGCCCAAGTCACTCTGTGGATC 2520  
Db 2461 ATCCAGAGGTTCAGATATCTTCAAGGTTTCTTATTTCTGGGGCCCAAGTCACTCTGTGGATC 2520  
QY 2521 TTCTATGATACAGGATACAGGAACTGTTATGTTGCTCACCTGACCAAGATGAACAGGCG 2580

Db 2521 TTCTATGATACAGGATACACGGAAAGTTATATGGGTCAACCTGACCCAGTAACAGAAATGAACAGGCG 2580  
QY 2581 TATTATTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAGACCAATCGTTTA 2640  
Db 2581 TATTATCTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAGACCAATCGTTTA 2640  
QY 2641 CTGCTCTTACATGGTTTCTGGATGAGAATGTCCATTTTGGCATACACAGATATATTACTG 2700  
Db 2641 CTGCTCTTACATGGTTTCTGGATGAGAATGTCCATTTTGGCATACACAGATATATTACTG 2700  
QY 2701 AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC 2760  
Db 2701 AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC 2760  
QY 2761 AGCATAAGAGTTCTGTAATCGGAGAAACATTATGAACCTGCATCTTTTGCACCTACCTTCAA 2820  
Db 2761 AGCATAAGAGTTCTGTAATCGGAGAAACATTATGAACCTGCATCTTTTGCACCTACCTTCAA 2820  
QY 2821 GAAACCTTGGATCACGTTATGCTCTTAAAGTGCATATAATTTTGGACCTGTGTAGAAC 2880  
Db 2821 GAAACCTTGGATCACGTTATGCTCTTAAAGTGCATATAATTTTGGACCTGTGTAGAAC 2880  
QY 2881 TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACACAGA 2940  
Db 2881 TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACACAGA 2940  
QY 2941 ATTGATCATCATTTTGTATACCTGCGCATGTAACTACTCTCTGAAATAAATTTGGTG 3000  
Db 2941 ATTGATCATCATTTTGTATACCTGCGCATGTAACTACTCTCTGAAATAAATTTGGTG 3000  
QY 3001 CcATGAGGGGTCTACGGTTTGTGTAGTAATCTAATACCTTTAACCCCAACATGCTCAAAA 3060  
Db 3001 CGATGAGGGGTCTACGGTTTGTGTAGTAATCTAATACCTTTAACCCCAACATGCTCAAAA 3060  
QY 3061 TCAATGATACATATTTCTGAGAGACCCAGCAATACCAATGATTAATAAATAAATAA 3120  
Db 3061 TCAATGATACATATTTCTGAGAGACCCAGCAATACCAATGATTAATAAATAAATAA 3120

## RESULT 2

US-10-825-632-2  
; Sequence 2, Application US/10825632  
; Publication No. US20040191826A1  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: FCSB-100-Div. 1  
; CURRENT APPLICATION NUMBER: US/10/825, 632  
; PRIOR FILING DATE: 2004-04-15  
; PRIOR APPLICATION NUMBER: US 10/070,464  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU P05709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU P02762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 3120  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-825-632-2

Query Match 100.0%; Score 3120; DB 8; Length 3120;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG 60  
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG 60

QY 61 CGTTCCGCGCTGGTTGTACCGGCGCGCCCGGAGGAAGCACTGCAACACAGGACCG 120  
Db 61 CGTTCCGCGCTGGTTGTACCGGCGCGCCCGGAGGAAGCACTGCAACACAGGACCG 120  
QY 121 GAGTGAGGCGCGCGCAGCATGAAGCGCGCAGCGCCCGCTCCATAGCGCAGCTCGGAGCGG 180  
Db 121 GAGTGAGGCGCGCGCAGCATGAAGCGCGCAGCGCCCGCTCCATAGCGCAGCTCGGAGCGG 180  
QY 181 TCCGCGCGCGCGCGGAGGAAGGAAATGCAA CATGGCAGCAGCAATGGAACACAGAACAG 240  
Db 181 TCCGCGCGCGCGCGGAGGAAGGAAATGCAA CATGGCAGCAGCAATGGAACACAGAACAG 240  
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGGAGGAGAAATTTGAAATCAAGGATCGG 300  
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGGAGGAGAAATTTGAAATCAAGGATCGG 300  
QY 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTAATCTGAGGAGTCAAGCTTAAAGAGCTGCTT 360  
Db 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTAATCTGAGGAGTCAAGCTTAAAGAGCTGCTT 360  
QY 361 GCCGATACAGAAATATCATGATGGCTAAGGCGCACCATGATGATTTTCATGTTT 420  
Db 361 GCCGATACAGAAATATCATGATGGCTAAGGCGCACCATGATGATTTTCATGTTT 420  
QY 421 GTGAGAGGAATGATCCAGATGGACCTCAITTCAGACAGAAATCTATTTACCTTGCCATGCT 480  
Db 421 GTGAGAGGAATGATCCAGATGGACCTCAITTCAGACAGAAATCTATTTACCTTGCCATGCT 480  
QY 481 GTGAGAGGAATGATCCAGATGGACCTCAITTCAGACAGAAATCTATTTACCTTGCCATGCT 480  
Db 481 GTGAGAGGAATGATCCAGATGGACCTCAITTCAGACAGAAATCTATTTACCTTGCCATGCT 480  
QY 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTGGATCTTTTTCAGGCAACACATGCACTAT 600  
Db 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTGGATCTTTTTCAGGCAACACATGCACTAT 600  
QY 601 GGAATGTTATCTCCAGAGGAAGAACTATTTAAGAGGAAGAAACGCAATGGAACAGTCGGA 660  
Db 601 GGAATGTTATCTCCAGAGGAAGAACTATTTAAGAGGAAGAAACGCAATGGAACAGTCGGA 660  
QY 661 ATTGCTCTTACGATTTACCAAGGAAGTGGAACTATTTCTGTTTCAAGCCGGTAGTGA 720  
Db 661 ATTGCTCTTACGATTTACCAAGGAAGTGGAACTATTTCTGTTTCAAGCCGGTAGTGA 720  
QY 721 ATTTATCACTGTAAGATGGAGGCCACAGGATTTTACGCAACAACTTTTAAGGCCCAAT 780  
Db 721 ATTTATCACTGTAAGATGGAGGCCACAGGATTTTACGCAACAACTTTTAAGGCCCAAT 780  
QY 781 CTAGTGGAACCTAGTTCTCCAA CATACGATGGATCCAAATTTATGCCCGCTGATCA 840  
Db 781 CTAGTGGAACCTAGTTCTCCAA CATACGATGGATCCAAATTTATGCCCGCTGATCA 840  
QY 841 GACTGGATTTGCTTTTATACATAGCAAGATTTTGGATCTATCAACATGCTACCCAGAGAA 900  
Db 841 GACTGGATTTGCTTTTATACATAGCAAGATTTTGGATCTATCAACATGCTACCCAGAGAA 900  
QY 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960  
Db 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA 960  
QY 961 GCTGGAGTCTGCTACCTTTGTTCTTCAAGAGAAATTTGATAGATATTTGCGCTATTGGTGG 1020  
Db 961 GCTGGAGTCTGCTACCTTTGTTCTTCAAGAGAAATTTGATAGATATTTGCGCTATTGGTGG 1020  
QY 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATCTCTATATGAAGA 1080  
Db 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATCTCTATATGAAGA 1080  
QY 1081 AATGATGAATCTGAGGTGGAATTTATTCATGTTTACATCCCTTATGTTGGAACCAAGGAGG 1140  
Db 1081 AATGATGAATCTGAGGTGGAATTTATTCATGTTTACATCCCTTATGTTGGAACCAAGGAGG 1140

QY	1141	GCAGATTCATTCCGTTATCCCTAAACAGGTACAGCAATCCCTAAAGTCACCTTTTAAGATG	1200
Db	1141	GCAGATTCATTCCGTTATCCCTAAACAGGTACAGCAATCCCTAAAGTCACCTTTTAAGATG	1200
QY	1201	TCAGAAATAATAGTATGATGCTGAAGAGGATCATAGATGTCATAGATAAGAACTAAAT	1260
Db	1201	TCAGAAATAATAGTATGATGCTGAAGAGGATCATAGATGTCATAGATAAGAACTAAAT	1260
QY	1261	CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTGGCAGAGCTGGATGACCTCCT	1320
Db	1261	CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTGGCAGAGCTGGATGACCTCCT	1320
QY	1321	GAGGAAAATATGCTTGGTCCATCCCTACTAGATCGCTCCGACACTCGCCTACAGATAGTG	1380
Db	1321	GAGGAAAATATGCTTGGTCCATCCCTACTAGATCGCTCCGACACTCGCCTACAGATAGTG	1380
QY	1381	TTGATCTCACTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Db	1381	TTGATCTCACTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
QY	1441	ATTCAGTCAGTCCGCTGATTCCTGACGCCCACTAATATCTATGAGAAACACAGACATC	1500
Db	1441	ATTCAGTCAGTCCGCTGATTCCTGACGCCCACTAATATCTATGAGAAACACAGACATC	1500
QY	1501	TGGAATAATATCCATGATCATCTTTCATGTTTCCCAAGTCAGAGAGGAATTCAG	1560
Db	1501	TGGAATAATATCCATGATCATCTTTCATGTTTCCCAAGTCAGAGAGGAATTCAG	1560
QY	1561	TTTATTTTGGCTCTGAAATCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAAATCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT	1620
QY	1621	TTAAAGGAAAGCAAAATATAACGATCCAGTGTGGCTGCCCTGCCCAAGTGATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAAAATATAACGATCCAGTGTGGCTGCCCTGCCCAAGTGATTTCAAG	1680
QY	1681	TGTCCTATCAAAGAGGAGATAGCAATTTACCAGTGTGGTGAATGGGAAGTTCTTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAAGAGGAGATAGCAATTTACCAGTGTGGTGAATGGGAAGTTCTTTGGCCGGCAT	1740
QY	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAAGCGCTGGTATATTTTGAAGGCACCAAGAGAC	1800
Db	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAAGCGCTGGTATATTTTGAAGGCACCAAGAGAC	1800
QY	1801	TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTCTGGAGAGGTCACAAGG	1860
Db	1801	TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTCTGGAGAGGTCACAAGG	1860
QY	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGCTCTTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGCTCTTTTATA	1920
QY	1921	AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTACAGCTATCAAGTCCT	1980
Db	1921	AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTACAGCTATCAAGTCCT	1980
QY	1981	GAAGATGACCCAACTGCAAAAACAAAGGAATTTGGGCCACCATTTTGGATTCAGCAGGT	2040
Db	1981	GAAGATGACCCAACTGCAAAAACAAAGGAATTTGGGCCACCATTTTGGATTCAGCAGGT	2040
QY	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
Db	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
QY	2101	TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG	2160
Db	2101	TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTG	2160
QY	2161	CTGTTCATATATGGTGGTCTCAGGTGTCAGTTGGTGAATTAATCGGTTTAAAGAGTCAAG	2220
Db	2161	CTGTTCATATATGGTGGTCTCAGGTGTCAGTTGGTGAATTAATCGGTTTAAAGAGTCAAG	2220
QY	2221	TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGTTAGTATAGACAAACAGG	2280

Db	2221	TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGTTAGTATAGACAAACAGG	2280
QY	2281	GGATCTCTGTCCACGAGGGCTTAAATTTTGAAGGGCCCTTTAAATATAAAATGGGTCAATA	2340
Db	2281	GGATCTCTGTCCACGAGGGCTTAAATTTTGAAGGGCCCTTTAAATATAAAATGGGTCAATA	2340
QY	2341	GAATTTCAACATCAGGTGGAGGAACCTCAATATCTAGCTTCTCGATATGATTTTCATTGAC	2400
Db	2341	GAATTTCAACATCAGGTGGAGGAACCTCAATATCTAGCTTCTCGATATGATTTTCATTGAC	2400
QY	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAGATACCTCTCCCTGATGGCATTA	2460
Db	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAGATACCTCTCCCTGATGGCATTA	2460
QY	2461	ATGCAGAGGTCAGATATCTTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCCTCTGTGGATC	2520
Db	2461	ATGCAGAGGTCAGATATCTTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCCTCTGTGGATC	2520
QY	2521	TTCTATGATACAGGATACAGGAACCTTATATGGGTCAACCCTGACCAAGATGAACAGGCG	2580
Db	2521	TTCTATGATACAGGATACAGGAACCTTATATGGGTCAACCCTGACCAAGATGAACAGGCG	2580
QY	2581	TATTTACTTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTA	2640
Db	2581	TATTTACTTAGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTA	2640
QY	2641	CTGCTCTTACATGGTTTCCCTGGATGAGAAATGTCATTTTGGACATACCAAGTATATTACTG	2700
Db	2641	CTGCTCTTACATGGTTTCCCTGGATGAGAAATGTCATTTTGGACATACCAAGTATATTACTG	2700
QY	2701	AGTTTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACAC	2760
Db	2701	AGTTTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACAC	2760
QY	2761	AGCATAAGAGTCTCTGAATCGGAGAACATTAATGAATCTTTTGCACCTACTCTTCAA	2820
Db	2761	AGCATAAGAGTCTCTGAATCGGAGAACATTAATGAATCTTTTGCACCTACTCTTCAA	2820
QY	2821	GAATAACCTTGGATCAGCATATTGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAAC	2880
Db	2821	GAATAACCTTGGATCAGCATATTGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAAC	2880
QY	2881	TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA	2940
Db	2881	TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA	2940
QY	2941	ATTGATCATCACATTTTGTATACCTGCATGTAACATCTACTCTGAAAAATAATGTGGTG	3000
Db	2941	ATTGATCATCACATTTTGTATACCTGCATGTAACATCTACTCTGAAAAATAATGTGGTG	3000
QY	3001	CCATGCAAGGGCTCTACGGTTTGGTGGTAACTAATACCTTAACCCACATGCTCAAAA	3060
Db	3001	CCATGCAAGGGCTCTACGGTTTGGTGGTAACTAATACCTTAACCCACATGCTCAAAA	3060
QY	3061	TCAAATGATACATATTCCTGAGAGACCCAGCAATACCATAAGATTTACTTAAAAA	3120
Db	3061	TCAAATGATACATATTCCTGAGAGACCCAGCAATACCATAAGATTTACTTAAAAA	3120

## RESULT 3

US-10-311-035-30  
; Sequence 30, Application US/10311035  
; Publication No. US20040023243A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: YUE, Henry  
; APPLICANT: ELLIOTT, Vicki  
; APPLICANT: GANDHI, Ameena R.  
; APPLICANT: LAL, Preeti  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: DELEGANE, Angelo M.



Db	1562	TCTGAATGCAAAACAGGTTCCGTCATATTTATACAAATTCATCTATTTTAAAGGAAAGC	1621
Qy	1633	AAATATAAACGATCCAGTGGTGGGCTCCCTGCTCCAAGTGATTTCAAAGTGTCTTACAA	1692
Db	1622	AAATATAAACGATCCAGTGGTGGGCTCCCTGCTCCAAGTGATTTCAAAGTGTCTTACAA	1681
Qy	1693	GAGGAGTAGCAATTAACAGTGGTGGAATGGGAAGTCTCTGGCCGGGCATGGATCTAAATATC	1752
Db	1682	GAGGAGTAGCAATTAACAGTGGTGGAATGGGAAGTCTCTGGCCGGGCATGGATCTAAATATC	1741
Qy	1753	CAAGTTGATGAAGTCAGAAAGCTGGTATATTTTTGAAGGCACCAAAAGATCCCCCTTTAGAG	1812
Db	1742	CAAGTTGATGAAGTCAGAAAGCTGGTATATTTTTGAAGGCACCAAAAGATCCCCCTTTAGAG	1801
Qy	1813	CATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGTGCACAAAGCTGACTGACCGT	1872
Db	1802	CATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGTGCACAAAGCTGACTGACCGT	1861
Qy	1873	GGCTACTCACATCTTCTGCTGCATCAGTCAGCATCTGTGACTCTCTTTATAGTAAAGTATAGT	1932
Db	1862	GGCTACTCACATCTTCTGCTGCATCAGTCAGCATCTGTGACTCTCTTTATAGTAAAGTATAGT	1921
Qy	1933	AACCCAGAGATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGCTCCTGAAGATGACCCA	1992
Db	1922	AACCCAGAGATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGCTCCTGAAGATGACCCA	1981
Qy	1993	ACTTCGAAAAAACHAGGAATTTTGGGCCACATTTTGGATTCAGCAGGTCTCTTCCCTGAC	2052
Db	1982	ACTTCGAAAAAACHAGGAATTTTGGGCCACATTTTGGATTCAGCAGGTCTCTTCCCTGAC	2041
Qy	2053	TATACTCTCCGAAAAATTTCTCTTTTGGAAAGTACTACGTGAATTCATCTGATGGGATG	2112
Db	2042	TATACTCTCCGAAAAATTTCTCTTTTGGAAAGTACTACGTGAATTCATCTGATGGGATG	2101
Qy	2113	CTCTACAAGCCCTCATGATCTACAGCCTCGGAAAGAAATATCCTTACTGTGCTCTTCATATAT	2172
Db	2102	CTCTACAAGCCCTCATGATCTACAGCCTCGGAAAGAAATATCCTTACTGTGCTCTTCATATAT	2161
Qy	2173	GGTGTGCTCTCAGGTGCAATGTTGGTGAATAATCCGGTTTAAAGGAGTCAAGTATTTCCGCTTG	2232
Db	2162	GGTGTGCTCTCAGGTGCAATGTTGGTGAATAATCCGGTTTAAAGGAGTCAAGTATTTCCGCTTG	2221
Qy	2233	AATAACCTTAGGCCTCTCTAGGTTATGTGTTGTAGTGATAGACAAACAGGGGATCCTGTTCAC	2292
Db	2222	AATAACCTTAGGCCTCTCTAGGTTATGTGTTGTAGTGATAGACAAACAGGGGATCCTGTTCAC	2281
Qy	2293	CGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAATAGAAATTTGACGAT	2352
Db	2282	CGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAATAGAAATTTGACGAT	2341
Qy	2353	CAGGTGGAAGCATCCCAATATCTAGCTCTCCGATATGATTTCAATTGACTTAGATCCGCTG	2412
Db	2342	CAGGTGGAAGCATCCCAATATCTAGCTCTCCGATATGATTTCAATTGACTTAGATCCGCTG	2401
Qy	2413	GGCATCCAGCGTGGTCCCTATGGAGGATCACTCTCCCTGATGGGCAATTAATGCAGAGGTCA	2472
Db	2402	GGCATCCAGCGTGGTCCCTATGGAGGATCACTCTCCCTGATGGGCAATTAATGCAGAGGTCA	2461
Qy	2473	GATATCTTCAAGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGTGTGAATCTTCTATGATACA	2532
Db	2462	GATATCTTCAAGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGTGTGAATCTTCTATGATACA	2521
Qy	2533	GGATACACGGAAACGTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGA	2592
Db	2522	GGATACACGGAAACGTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGA	2581
Qy	2593	TCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTACAT	2652
Db	2582	TCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTACAT	2641
Qy	2653	GGTTTCTCTGGATGAGAAATGTCCATTTTGTGACATACAGCATATATTACTGAGTTTTTTAGTG	2712
Db	2642	GGTTTCTCTGGATGAGAAATGTCCATTTTGTGACATACAGCATATATTACTGAGTTTTTTAGTG	2701

RESULT 4

US-10-170-789-37

Sequence 37, Application US/10170789

Publication NO. US20030180930A1

GENERAL INFORMATION:

APPLICANT: Rachel E. Meyers

APPLICANT: Olandt, Peter J.

APPLICANT: Kapeller-Libermann, Rosana

APPLICANT: Curtis, Rory A. J.

APPLICANT: Williamson, Mark

APPLICANT: Weich, Nadine

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE, AND PROTEASE FAMILY MEMBERS AND USES THEREOF

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE, AND PROTEASE FAMILY MEMBERS AND USES THEREOF

FILE REFERENCE: 10448-191001

CURRENT APPLICATION NUMBER: US/10/170,789

CURRENT FILING DATE: 2002-06-13

PRIOR APPLICATION NUMBER: US 09/797,039

PRIOR FILING DATE: 2001-02-28

PRIOR APPLICATION NUMBER: PCT/US01/06525

PRIOR FILING DATE: 2001-02-28

PRIOR APPLICATION NUMBER: US 60/186,061

PRIOR FILING DATE: 2000-02-29

PRIOR APPLICATION NUMBER: US 09/882,166

PRIOR FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: PCT/US01/19269

PRIOR FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: US 60/212,078

PRIOR FILING DATE: 2000-06-15

PRIOR APPLICATION NUMBER: US 09/934,406

PRIOR FILING DATE: 2001-08-21

PRIOR APPLICATION NUMBER: PCT/US01/36052

PRIOR FILING DATE: 2001-08-21

PRIOR APPLICATION NUMBER: US 60/226,740

PRIOR FILING DATE: 2000-08-21

PRIOR APPLICATION NUMBER: US 09/861,801

PRIOR FILING DATE: 2001-05-21

PRIOR APPLICATION NUMBER: PCT/US01/16549

PRIOR FILING DATE: 2001-05-21

PRIOR APPLICATION NUMBER: US 60/205,508

PRIOR FILING DATE: 2000-05-19

PRIOR APPLICATION NUMBER: US 09/801,267

PRIOR FILING DATE: 2001-03-06

PRIOR APPLICATION NUMBER: PCT/US01/07138



```

; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(2874)
US-10-170-789-37

Query Match          94.8%; Score 2957; DB 6; Length 3143;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 GCCTCCGAGGCGAAGCGCGTCTACTGCGCGCGTCTTCTTAGTGCAGCGGTTCGCGCG 70
DB 26 GCCTCCGAGGCGAAGCGCGTCTACTGCGCGCGTCTTCTTAGTGCAGCGGTTCGCGCG 85
QY 71 CTGGTGTCAACCGCGCGCGCGCGAGGAGCCACTGCAACAGGACCGAGTGGAGGC 130
DB 86 CTGGTGTCAACCGCGCGCGCGCGAGGAGCCACTCAACAGGACCGAGTGGAGGC 145
QY 131 GGCAGCATGAGCGCGCGCGCGCGCGCTCCATAGGCGACGTGCGGACGCTCCGCGCGG 190
DB 146 GGCAGCATGAGCGCGCGCGCGCGCGCTCCATAGGCGACGTGCGGACGCTCCGCGCGG 205
QY 191 GCGCGGGGAGGAAATGCAACATGCGAGCGCAATGGAACAGACAGCTGGGTGTG 250
DB 206 GCGCGGGGAGGAAATGCAACATGCGAGCGCAATGGAACAGACAGCTGGGTGTG 265
QY 251 AGATATTTGAACTGCGGACTGTGAGGAGAAATATTGAATCAAGGATCGGCTAAATTTGG 310
DB 266 AGATATTTGAACTGCGGACTGTGAGGAGAAATATTGAATCAAGGATCGGCTAAATTTGG 325
QY 311 AdCCTTTTATGTTAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTTCCCGATACCA 370
DB 326 AdCCTTTTATGTTAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTTCCCGATACCA 385
QY 371 GAAATATCATGGCTACATGATGCTTAAGGCAACCATGATTTCTATGTTTGTGAAGGGA 430
DB 386 GAAATATCATGGCTACATGATGCTTAAGGCAACCATGATTTCTATGTTTGTGAAGGGA 445
QY 431 ATGATCCAGATGGACCTCATATTCAGACAGAAATCTATTACCTTGCATGTCTGGTGAACA 490
DB 446 ATGATCCAGATGGACCTCATATTCAGACAGAAATCTATTACCTTGCATGTCTGGTGAACA 505
QY 491 GAGAAATACACTGTTTATTCTGAAATTTCCCAAACTATCAATAGACGACGCTTAA 550
```

```

DB 506 GAGAAATACACTGTTTATTCTGAAATTTCCCAAACTATCAATAGACGACGCTTAA 565
QY 551 TGCTCTCTTGGAGAGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTATGGAATGATT 610
DB 566 TGCTCTCTTGGAGAGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTATGGAATGATT 625
QY 611 CTCGAGAAAGAACTATTAAAGAAAGAAACGCAATTTGGAACAGCTCGGAATTTGCTTCTT 670
DB 626 CTCGAGAAAGAACTATTAAAGAAAGAAACGCAATTTGGAACAGCTCGGAATTTGCTTCTT 685
QY 671 ACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAG 730
DB 686 ACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAG 745
QY 731 TAAAGATGGAGGGCCACCAAGGATTTACGCAACCACTTTAAGGCCCAATCTAGTGGAA 790
DB 746 TAAAGATGGAGGGCCACCAAGGATTTACGCAACCACTTTAAGGCCCAATCTAGTGGAA 805
QY 791 CTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCCCTGATCCAGACTGGATTG 850
DB 806 CTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCCCTGATCCAGACTGGATTG 865
QY 851 CTTTATACATACGACGATTTTGGATATCTAAACATGTAACCAAGAAAGAGGAC 910
DB 866 CTTTATACATACGACGATTTTGGATATCTAAACATGTAACCAAGAAAGAGGAC 925
QY 911 TCACCTATGTGCAATAGTAGCCACATCGAAGAGATGCCAGATCAGCTGGAGTCG 970
DB 926 TCACCTATGTGCAATAGTAGCCACATCGAAGAGATGCCAGATCAGCTGGAGTCG 985
QY 971 CTACCTTTGTTCTCAAGGAAGATTTGATAGATATTTCTGGCTATTGGTGTCCAAAG 1030
DB 986 CTACCTTTGTTCTCAAGGAAGATTTGATAGATATTTCTGGCTATTGGTGTCCAAAG 1045
QY 1031 CTGAAACAACTCCAGTGGTGTAAATTTCTAGATTTCTATATGAAGAAATGATGAAT 1090
DB 1046 CTGAAACAACTCCAGTGGTGTAAATTTCTAGATTTCTATATGAAGAAATGATGAAT 1105
QY 1091 CTGAGGTGGAATTTATTTACATTCCTCTATGTTGAAACAGGAGGCGAGTTTCAT 1150
DB 1106 CTGAGGTGGAATTTATTTACATTCCTCTATGTTGAAACAGGAGGCGAGATTTCAT 1165
QY 1151 TCCGTATCTCTAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGTGTGAGAAATA 1210
DB 1166 TCCGTATCTCTAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGTGTGAGAAATA 1225
QY 1211 TGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACATAATTTCAACCTTTG 1270
DB 1226 TGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACATAATTTCAACCTTTG 1285
QY 1271 AGATTCTATTGGAAGGATTTGAATATTTCCAGAGCTGGATGGACTCCTGAGGGAAT 1330
DB 1286 AGATTCTATTGGAAGGATTTGAATATTTCCAGAGCTGGATGGACTCCTGAGGGAAT 1345
QY 1331 ATGCTTGGTCCATCTCTAATAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAC 1390
DB 1346 ATGCTTGGTCCATCTCTAATAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAC 1405
QY 1391 CTGAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGCGAGACTCATTGAGTCCAG 1450
DB 1406 CTGAATTTATTTATCCAGTAGAAGATGATTTATGGAAGGCGAGACTCATTGAGTCCAG 1465
QY 1451 TGCCTGATTCTGTGACGCCACTAATTTATCTATGAGAAACACAGACATCTGGATAAATA 1510
DB 1466 TGCCTGATTCTGTGACGCCACTAATTTATCTATGAGAAACACAGACATCTGGATAAATA 1525
QY 1511 TCCATGACATCTTTTCATGTTTTTTTCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGG 1570
DB 1526 TCCATGACATCTTTTCATGTTTTTTTCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGG 1585
QY 1571 CCTCTGAATGCMAAACAGGTTTCGGTCATTTATACAAATTTACATCTATTTTAAAGGAAA 1630
```

Db 1586 CCTCTGAATGCAAAACAGGTTTCGGTCAATTATACAAAATTACATCTATTATTAAGGAAA 1645  
Qy 1631 GCRAATATAACGATCCAGTGGTGGCTGCTCTCCAAAGTGAATTTCAAGTGTCTATCA 1690  
Db 1646 GCAATATAACGATCCAGTGGTGGCTGCTCTCCAAAGTGAATTTCAAGTGTCTATCA 1705  
Qy 1691 AAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATA 1750  
Db 1706 AAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATA 1765  
Qy 1751 TCCAAGTTGATGAAGTGAAGGCTGGTATATTTTGAAGGCCAAAGACTCCCTTTAG 1810  
Db 1766 TCCAAGTTGATGAAGTGAAGGCTGGTATATTTTGAAGGCCAAAGACTCCCTTTAG 1825  
Qy 1811 AGCATCACTGTAGTCAAGTACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACC 1870  
Db 1826 AGCATCACTGTAGTCAAGTACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACC 1885  
Qy 1871 GTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATA 1930  
Db 1886 GTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATA 1945  
Qy 1931 GTAACCAAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCTGAAGTACCC 1990  
Db 1946 GTAACCAAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCTGAAGTACCC 2005  
Qy 1991 CAACCTTCAAAACAAAGGAATTTTGGGCCACCAATTTGGATTCAGCAGGTCTCTCTG 2050  
Db 2006 CAACCTTCAAAACAAAGGAATTTTGGGCCACCAATTTGGATTCAGCAGGTCTCTCTG 2065  
Qy 2051 ACTATAGTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATGTATGGGA 2110  
Db 2066 ACTATAGTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATGTATGGGA 2125  
Qy 2111 TGCTCTACAAGCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTGATAT 2170  
Db 2126 TGCTCTACAAGCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTGATAT 2185  
Qy 2171 ATGCTGTCTCAGGTGCAATGCTGGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCT 2230  
Db 2186 ATGCTGTCTCAGGTGCAATGCTGGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCT 2245  
Qy 2231 TGAATACCTTAGCTCTTAGGTTATGTGTTGTAGTGATAGACAAAGGGGATCTCTGTC 2290  
Db 2246 TGAATACCTTAGCTCTTAGGTTATGTGTTGTAGTGATAGACAAAGGGGATCTCTGTC 2305  
Qy 2291 ACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATAGAAATGACG 2350  
Db 2306 ACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATAGAAATGACG 2365  
Qy 2351 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTG 2410  
Db 2366 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTG 2425  
Qy 2411 TGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAGGTT 2470  
Db 2426 TGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAGGTT 2485  
Qy 2471 CAGATATCTTCAGGTTGCTATTTCTGGGGCCCGAGTCATCTGTGGATCTTCTATGATA 2530  
Db 2486 CAGATATCTTCAGGTTGCTATTTCTGGGGCCCGAGTCATCTGTGGATCTTCTATGATA 2545  
Qy 2531 CAGATACAGGACGTTATATGGGTCAACCTGACAGATGAACAGGGCTATTACTTAG 2590  
Db 2546 CAGATACAGGACGTTATATGGGTCAACCTGACAGATGAACAGGGCTATTACTTAG 2605  
Qy 2591 GATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTTTAC 2650  
Db 2606 GATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTTTAC 2665  
Qy 2651 ATGGTTTCTTGGATGAGAAATGTCATTTTGACATACAGPATATTAATCAGTATTTTATAG 2710  
Db 2666 ATGGTTTCTTGGATGAGAAATGTCATTTTGACATACAGPATATTAATCAGTATTTTATAG 2725

## RESULT 5

US-09-976-674-12  
; Sequence 12, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 4829  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-12

Query Match 92.6%; Score 2889; DB 3; Length 4829;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 1 AAGTGTCTAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCCCGCTGCTTCTTAGTGCCG 60  
Db 1 AAGTGTCTAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCCCGCTGCTTCTTAGTGCCG 60  
Qy 61 CGTTGCGCGCTGGGTGTGTCACCGCGCGCGCGCGAGGAAGCCACTGCAACCAAGSACCG 120  
Db 61 CGTTGCGCGCGCTGGGTGTGTCACCGCGCGCGCGCGAGGAAGCCACTGCAACCAAGSACCG 120  
Qy 121 GAGTGAAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTGCGGACCG 180  
Db 121 GAGTGAAGCGCGCGCAGCATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTGCGGACCG 180  
Qy 181 TCCGGCGCGCGCGCGGGAAGAAAATGCAACATGGCAGCAGCAATGGAAAACAGAACAG 240  
Db 181 TCCGGCGCGCGCGCGGGAAGAAAATGCAACATGGCAGCAGCAATGGAAAACAGAACAG 240

QY 241 CTGGGTGTGAGATATTTTGAACCTCGCGACTGTGAGGAGAAATATTGAATACAGGATCGG 300  
Db 241 CTGGGTGTGAGATATTTTGAACCTCGCGACTGTGAGGAGAAATATTGAATACAGGATCGG 300  
QY 301 CCTAAATTTGAGGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTTAAAAAGCTGCTT 360  
Db 301 CCTAAATTTGAGGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTTAAAAAGCTGCTT 360  
QY 361 GCGCATACCCAGAAAAATCATGGCTACATGATGGCTTAAGGCACACATGATTTTCATGTTT 420  
Db 361 GCCGATACCCAGAAAAATCATGGCTACATGATGGCTTAAGGCACACATGATTTTCATGTTT 420  
QY 421 GTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGAACTATTATTCCTTTGCCATGTCT 480  
Db 421 GTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGAACTATTATTCCTTTGCCATGTCT 480  
QY 481 GGTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540  
Db 481 GGTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCA 540  
QY 541 GCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
Db 541 GCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
QY 601 GGAATGTTTCTCGAGAGAGAACTATTAAAGAAAGAAACCGCATTTGGAACAGTCGGA 660  
Db 601 GGAATGTTTCTCGAGAGAGAACTATTAAAGAAAGAAACCGCATTTGGAACAGTCGGA 660  
QY 661 ATTGCTTCTTACGATATATCACCAAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA 720  
Db 661 ATTGCTTCTTACGATATATCACCAAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA 720  
QY 721 ATTTATCACTGTAAGATGAGGGCCACAAAGATTTAGCAACAACTTTTAAAGCCCAAT 780  
Db 721 ATTTATCACTGTAAGATGAGGGCCACAAAGATTTAGCAACAACTTTTAAAGCCCAAT 780  
QY 781 CTAGTGGAAACTAGTTGTCTCCAAACATACGGAATGGAATCCAAATATTATGCCCGCTGATCCA 840  
Db 781 CTAGTGGAAACTAGTTGTCTCCAAACATACGGAATGGAATCCAAATATTATGCCCGCTGATCCA 840  
QY 841 GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACAGAGAA 900  
Db 841 GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACAGAGAA 900  
QY 901 GAAAGGAGACTCACTATGTGCAATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960  
Db 901 GAAAGGAGACTCACTATGTGCAATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960  
QY 961 GCTGGAGTCGTAACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020  
Db 961 GCTGGAGTCGTAACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020  
QY 1021 TGTCCAAAGCTGAAACAACTCCAGTGTGTGTAATAATTTCTAGAAATTTCTATAGAAGAA 1080  
Db 1021 TGTCCAAAGCTGAAACAACTCCAGTGTGTGTAATAATTTCTAGAAATTTCTATAGAAGAA 1080  
QY 1081 AATGATGAATCTGAGGTGGAATTTATTCATGTTTACATCCCTCATGTTGGAACAAAGAGG 1140  
Db 1081 AATGATGAATCTGAGGTGGAATTTATTCATGTTTACATCCCTCATGTTGGAACAAAGAGG 1140  
QY 1141 GAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCTTAAAGTCATCTTTTAAAGATG 1200  
Db 1141 GAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCTTAAAGTCATCTTTTAAAGATG 1200  
QY 1201 TCAGAAATTAATGATGCTGGAAGGAAGATCATAGATGCTCATAGATAAGGAACATAAT 1260  
Db 1201 TCAGAAATTAATGATGCTGGAAGGAAGATCATAGATGCTCATAGATAAGGAACATAAT 1260  
QY 1261 CAACCTTTGAGATTTCTATTTGAAGGATTTGAATATTTTGGCAGAGCTGGATGACTCCT 1320  
Db 1261 CAACCTTTGAGATTTCTATTTGAAGGATTTGAATATTTTGGCAGAGCTGGATGACTCCT 1320  
QY 1321 GAGGAAAAATATGCTTGGTCCATCTTACTAGATCGCTCCGAGCTCGCTACAGATAGTG 1380

Db 1321 GAGGAAAAATATGCTTGGTCCATCTTACTAGATCGCTCCGAGCTCGCTACAGATAGTG 1380  
QY 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGACAGAGACTC 1440  
Db 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGACAGAGACTC 1440  
QY 1441 ATTGAGTCAGTCTGATTTCTGTGAGCCCACTAAATTTCTATGGAAGAAACAAACAGACATC 1500  
Db 1441 ATTGAGTCAGTCTGATTTCTGTGAGCCCACTAAATTTCTATGGAAGAAACAAACAGACATC 1500  
QY 1501 TGGATAAATATCCAAGACATCTTTTTCATGTTTTTCCCAAGGTCACGAAGAGGAAATTCGAG 1560  
Db 1501 TGGATAAATATCCAAGACATCTTTTTCATGTTTTTCCCAAGGTCACGAAGAGGAAATTCGAG 1560  
QY 1561 TTTATTTTTCCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTATT 1620  
Db 1561 TTTATTTTTCCTCTGAAATGCAAAACAGGTTTCCGTCATTTTATACAAAATTTACATCTATT 1620  
QY 1621 TTAAGGAAAGCAAAATATAAAGATCCAGTGTGTGGCTGCCCTGCCAAGTGAATTTTCAG 1680  
Db 1621 TTAAGGAAAGCAAAATATAAAGATCCAGTGTGTGGCTGCCCTGCCAAGTGAATTTTCAG 1680  
QY 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGTGTGAATGGAAGTTCTTTGGCCGGCAT 1740  
Db 1681 TGTCTTATCAAGAGGAGATAGCAATTTACAGTGTGTGAATGGAAGTTCTTTGGCCGGCAT 1740  
QY 1741 GGATCTAATATCCAAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800  
Db 1741 GGATCTAATATCCAAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800  
QY 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAG 1860  
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAG 1860  
QY 1861 CTGATGACCGTGGCTACTCAGATTTCTGTGATCAGTCAGCAGCTGTGACTTCTTTTATA 1920  
Db 1861 CTGATGACCGTGGCTACTCAGATTTCTGTGATCAGTCAGCAGCTGTGACTTCTTTTATA 1920  
QY 1921 AGTAAGTATAGTAACCAAGAGATCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCT 1980  
Db 1921 AGTAAGTATAGTAACCAAGAGATCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCT 1980  
QY 1981 GAAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGCCCACTTTTGGATTCAGCAGGT 2040  
Db 1981 GAAGATGACCCAACTTGCAGAAACAAAGGAATTTTGGCCCACTTTTGGATTCAGCAGGT 2040  
QY 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTACA 2100  
Db 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTACA 2100  
QY 2101 TTGATGAGGATGCTTACAGCCCTCATGATCTACAGCCCTGGAAGAAATATCTCTACTGTG 2160  
Db 2101 TTGATGAGGATGCTTACAGCCCTCATGATCTACAGCCCTGGAAGAAATATCTCTACTGTG 2160  
QY 2161 CTGTTCTATATATGAGTGG--TCTCTCAGGTGAGTGGTGAATTAATCGGTTTAAAGAGTCA 2218  
Db 2161 CTGTTCTATATATGAGTGGTCTCTCTCAGGTGAGTGGTGAATTAATCGGTTTAAAGAGTCA 2220  
QY 2219 AGTATTTTCCGTTGAAATACCTAGCCCTCTAGGTTTATGTTGTTCTAGTATAGACAACA 2278  
Db 2219 AGTATTTTCCGTTGAAATACCTAGCCCTCTAGGTTTATGTTGTTCTAGTATAGACAACA 2280  
QY 2279 GGGGATCCTGTCAACCGAGGCTTTAAATTTGAAGGGCCCTTTTAAATATAAATTTGGGTCAAA 2338  
Db 2279 GGGGATCCTGTCAACCGAGGCTTTAAATTTGAAGGGCCCTTTTAAATATAAATTTGGGTCAAA 2340  
QY 2339 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTG 2398  
Db 2339 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTG 2400  
QY 2399 ACTTAGATCGTGTGGGATCCACGGCTGTCTTATGAGGATACCTCTCCCTGTAGTGCAT 2458  
Db 2399 ACTTAGATCGTGTGGGATCCACGGCTGTCTTATGAGGATACCTCTCCCTGTAGTGCAT 2458

Db	2401	ACTTAGATCGCTGGGCATCCACGGCTGTGCTCTATATGGAGGATACCTCTCTCCCTGATGGCAT	2466
Qy	2459	TAAATGACAGAGTCAGATATCTTACAGGTTGCTATATGCTGGGGCCCAAGTCACTCTGTGGGA	2518
Db	2461	TAAATGAGAGGTCAGATATCTTACAGGTTGCTATATGCTGGGGCCCAAGTCACTCTGTGGGA	2520
Qy	2519	TCTTCTATGATACAGGATACACGGAAAGCTTATATAGGGTCACCTTCACACAGAAATGAACAGG	2578
Db	2521	TCTTCTATGATACAGGATACACGGAAAGCTTATATAGGGTCACCTTCACACAGAAATGAACAGG	2580
Qy	2579	GCTATTAATTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCTCTGAAACCAATCGTT	2638
Db	2581	GCTATTAATTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCTCTGAAACCAATCGTT	2640
Qy	2639	TACTGCTCTTACATGGTTTCTCTGGATGAGAAATCTCAATTTTGCACATACCAAGTATATTAC	2698
Db	2641	TACTGCTCTTACATGGTTTCTCTGGATGAGAAATCTCAATTTTGCACATACCAAGTATATTAC	2700
Qy	2699	TGAGTTTTTTTGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGAC	2758
Db	2701	TGAGTTTTTTTGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGAC	2760
Qy	2759	ACAGCATAAAGAGTCTCTGAATTCGGGAGAACATTTATGAACTGCACTCTTTTGCACACTACCTTC	2818
Db	2761	ACAGCATAAAGAGTCTCTGAATTCGGGAGAACATTTATGAACTGCACTCTTTTGCACACTACCTTC	2820
Qy	2819	AAGAAAAACCTTGGATCACGTTATGCTGCTCTTAAAGTGATATAATTTTGAACCTGTGTAGA	2878
Db	2821	AAGAAAAACCTTGGATCACGTTATGCTGCTCTTAAAGTGATATAATTTTGAACCTGTGTAGA	2880
Qy	2879	ACTCTCTGGTATACATCTGGCTATTATAACCAAAATGAGGAGGTTTAAATCAACAGAAAAACA	2938
Db	2881	ACTCTCTGGTATACATCTGGCTATTATAACCAAAATGAGGAGGTTTAAATCAACAGAAAAACA	2940
Qy	2939	GAAATTGATCATCACATTTTGTATCCTGCCATGTAAACATCTACTCTCGAAAAATAAATGTGG	2998
Db	2941	GAAATTGATCATCACATTTTGTATCCTGCCATGTAAACATCTACTCTCGAAAAATAAATGTGG	3000
Qy	2999	TGCCATGACGAGGGTCTACGGTTTTGTGTAGTAAATCTAATACCTTAAACCCCAATGCTCAA	3058
Db	3001	TGCCATGACGAGGGTCTACGGTTTTGTGTAGTAAATCTAATACCTTAAACCCCAATGCTCAA	3060
Qy	3059	AATCAAAATGATACATATTCTTGAGAGACCCAGCAATACCATAAAGAAATTTACTTAAAAAAA	3118
Db	3061	AATCAAAATGATACATATTCTTGAGAGACCCAGCAATACCATAAAGAAATTTACTTAAAAAAA	3120
Qy	3119	AA	3120
Db	3121	AA	3122

## RESULT 6

```

US-10-982-512-12
: Sequence 12, Application US/10982512
: Publication No. US20050059081A1
: GENERAL INFORMATION:
: APPLICANT: Qi, Steve
: APPLICANT: Akinsanya, Karen
: APPLICANT: Riviere, Pierre
: APPLICANT: Junien, Jean-Louis
: TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
: FILE REFERENCE: 70669
: CURRENT APPLICATION NUMBER: US/10/982,512
: CURRENT FILING DATE: 2004-11-05
: PRIOR APPLICATION NUMBER: US/09/976,674
: PRIOR FILING DATE: 2001-10-12
: PRIOR APPLICATION NUMBER: US 60/240,117
: PRIOR FILING DATE: 2000-10-12
: NUMBER OF SEQ. IDS: 1
: SOFTWARE: Patent in version 3.1
: SEQ ID NO 12
: LENGTH: 4829
: TYPE: DNA

```

Db	961	GCTGGAGTCGCTACTCCTTTGTCTCCAAAGAANAATTTGATAGATATCTTGGCCTATTGGTG	1021
Qy	1021	TGTCCAAAGAGCTGAACAACCTCCCAAGTGGTGTAAAAATTTCTTAGAATTTCTATATGAAGAA	1080
Db	1021	TGTCCAAAGAGCTGAACAACCTCCCAAGTGGTGTAAAAATTTCTTAGAATTTCTATATGAAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGGAAAATTAATTCATGTTTACATCCCCTATGTTGGAAACAAGGAGG	1140
Db	1081	AATGATGAATCTGAGGTGGAAAATTAATTCATGTTTACATCCCCTATGTTGGAAACAAGGAGG	1140
Qy	1141	GAGATTCATTCCTCGTTATCTTAAACAAGGTACAGCAAAATCCCTAAAGTCACTCTTTTAAAGATG	1200
Db	1141	GAGATTCATTCCTCGTTATCTTAAACAAGGTACAGCAAAATCCCTAAAGTCACTCTTTTAAAGATG	1200
Qy	1201	TCAGAAATAATGATTTGATCTGAAGGAGGATCATAGATGTCATAGATAAGGAACTAAAT	1260
Db	1201	TCAGAAATAATGATTTGATCTGAAGGAGGATCATAGATGTCATAGATAAGGAACTAAAT	1260
Qy	1261	CACCTTTTTGAGATTCATATTTGAAGGAGTTTGAATATATTTGCCAGAGCTGGATGCACCTCT	1320
Db	1261	CACCTTTTTGAGATTCATATTTGAAGGAGTTTGAATATATTTGCCAGAGCTGGATGCACCTCT	1320
Qy	1321	GAGGAAAAATATGCTTGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
Db	1321	GAGGAAAAATATGCTTGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
Qy	1381	TTGATCTCACCTGAAATTAATTTATCCCAAGTAAAGATGATGTTATGCGAAAGCCAGAGATC	1440
Db	1381	TTGATCTCACCTGAAATTAATTTATCCCAAGTAAAGATGATGTTATGCGAAAGCCAGAGATC	1440
Qy	1441	ATTGAGTCAGTCGCTGATCTGTGACGCGCACTAAATTAATCTATGAAGAAACAAGAGATC	1500
Db	1441	ATTGAGTCAGTCGCTGATCTGTGACGCGCACTAAATTAATCTATGAAGAAACAAGAGATC	1500
Qy	1501	TGATAAATATCCATGACATCTTTTCATGTTTTTCCCCCAAAGTCACGAGAGGAAATTTGAG	1560
Db	1501	TGATAAATATCCATGACATCTTTTCATGTTTTTCCCCCAAAGTCACGAGAGGAAATTTGAG	1560
Qy	1561	TTTATTTTTGCTCTGAAATGCAAAACAGGTTTCCGTCATTATATACAAAATTCATCTATT	1620
Db	1561	TTTATTTTTGCTCTGAAATGCAAAACAGGTTTCCGTCATTATATACAAAATTCATCTATT	1620
Qy	1621	TTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCCTGCTCCAAGTGATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGGCTGCCTGCTCCAAGTGATTTCAAG	1680
Qy	1681	TGTCCTATCAAAGAGGAGATAGCAATTAACAAGTGGTGAATGGGAAGTTCTTGGCGCGCAT	1740
Db	1681	TGTCCTATCAAAGAGGAGATAGCAATTAACAAGTGGTGAATGGGAAGTTCTTGGCGCGCAT	1740
Qy	1741	GDATCTAAATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
Db	1741	GDATCTAAATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
Qy	1801	TCCCTTTTAGAGCATCACTGTACGTAGTCAGTTTACGTAAATCCCTGGAGAGGTCACAAGG	1860
Db	1801	TCCCTTTTAGAGCATCACTGTACGTAGTCAGTTTACGTAAATCCCTGGAGAGGTCACAAGG	1860
Qy	1861	CTGACGTACCCTGGCTACTCATTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTATA	1920
Db	1861	CTGACGTACCCTGGCTACTCATTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTATA	1920
Qy	1921	ADTAAGTATAGTAACACAGAGAATCCACATGTGTGTCCCTTTTACAAGCTATCAAGTCCT	1980
Db	1921	ADTAAGTATAGTAACACAGAGAATCCACATGTGTGTCCCTTTTACAAGCTATCAAGTCCT	1980
Qy	1981	GAGATGACCCAACTTGC AAAACAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT	2040
Db	1981	GAGATGACCCAACTTGC AAAACAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGT	2040
Qy	2041	CCTCTTCCTGACTATACCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
Db	2041	CCTCTTCCTGACTATACCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACA	2100

Qy	2101	TTGTATGGGATGCTCTTCAAGACCTCATGATGATCTACAGCCTTGGAAAGAAATATCTCTACTGTG	2160
Db	2101	TTGTATGGGATGCTCTTCAAGACCTCATGATCTACAGCCTTGGAAAGAAATATCTCTACTGTG	2160
Qy	2161	CTGTTTCATATATATGGTG--TCCTCAGGTGCAGTTTGGTGAATTAATCGGTTTAAAGGAGTCA	2218
Db	2161	CTGTTTCATATATATGGTGCTCTCAGGTGCAGTTTGGTGAATTAATCGGTTTAAAGGAGTCA	2220
Qy	2219	AGTATTTCCGCTTGAATTAACCTTAGCCTCTCTAGGTTTATGTGTTCTAGTGTATAGACAACA	2278
Db	2221	AGTATTTCCGCTTGAATTAACCTTAGCCTCTCTAGGTTTATGTGTTCTAGTGTATAGACAACA	2280
Qy	2279	GGGGATCCTGTCTCACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATATAAAATGGGTCAAA	2338
Db	2281	GGGGATCCTGTCTCACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATATAAAATGGGTCAAA	2340
Qy	2339	TAGAAATTTGACGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATATGATTTCAATTG	2398
Db	2341	TAGAAATTTGACGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATATGATTTCAATTG	2400
Qy	2399	ACTTAGATCGTGTGGGCATCTCACGGCTGTCTTATGGAGGATACCTCTCCCTGATGGCAT	2458
Db	2401	ACTTAGATCGTGTGGGCATCTCACGGCTGTCTTATGGAGGATACCTCTCCCTGATGGCAT	2460
Qy	2459	TAATGACAGGTCAGATATCTTACGGGTGCTATTCCTGGGGCCCCAGTCACTCTGTGGA	2518
Db	2461	TAATGACAGGTCAGATATCTTACGGGTGCTATTCCTGGGGCCCCAGTCACTCTGTGGA	2520
Qy	2519	TCTTCTATGATACAGGATACACGGAACTTATATGGTTCACCTCACAGATCAACAGG	2578
Db	2521	TCTTCTATGATACAGGATACACGGAACTTATATGGTTCACCTCACAGATCAACAGG	2580
Qy	2579	GCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTT	2638
Db	2581	GCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTT	2640
Qy	2639	TACTGCTCTTACATGGTTTCTCGATCAGAAATGCTCAATTTTGCATATACCAGTATATTAC	2698
Db	2641	TACTGCTCTTACATGGTTTCTCGATCAGAAATGCTCAATTTTGCATATACCAGTATATTAC	2700
Qy	2699	TGAGTTTTTTTACTGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGAC	2758
Db	2701	TGAGTTTTTTTACTGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGAC	2760
Qy	2759	ACAGCATAGAGTTCCTGAATCGGGAGAACATTTAGAACTGGCATCTTTTGCACCTACCTTC	2818
Db	2761	ACAGCATAGAGTTCCTGAATCGGGAGAACATTTAGAACTGGCATCTTTTGCACCTACCTTC	2820
Qy	2819	AAGAAAACCTTGATCAGTATTCCTCTCTTAAAGTGATATAATTTTGACCTGTGTAGA	2878
Db	2821	AAGAAAACCTTGATCAGTATTCCTCTCTTAAAGTGATATAATTTTGACCTGTGTAGA	2880
Qy	2879	ACTCTCTGGTATACACTGGCTATTTTAAACAAATGAGGAGTTTAAATCAACAGAAAACACA	2938
Db	2881	ACTCTCTGGTATACACTGGCTATTTTAAACAAATGAGGAGTTTAAATCAACAGAAAACACA	2940
Qy	2939	GAATTGATCATCACTTTTGTATCTCGCATGTAAACATCTACTCTCTGAAAAATAATGTGG	2998
Db	2941	GAATTGATCATCACTTTTGTATCTCGCATGTAAACATCTACTCTCTGAAAAATAATGTGG	3000
Qy	2999	TGCCATGCAAGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCCAATGCTCAA	3058
Db	3001	TGCCATGCAAGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCCAATGCTCAA	3060
Qy	3059	AATCAAATGATACATATTTCTTGAGAGACCCAGCAATACCATTAAGAATTACTTAAAAAAA	3118
Db	3061	AATCAAATGATACATATTTCTTGAGAGACCCAGCAATACCATTAAGAATTACTTAAAAAAA	3120
Qy	3119	AA 3120	
Db	3121	AA 3122	

```
RESULT 7
; Sequence 39, Application US/10170789
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Willis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 1048-191001
; CURRENT FILING DATE: 2002-06-13
; PRIOR FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2000-02-29
; PRIOR FILING DATE: 2000-02-29
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2000-06-15
; PRIOR FILING DATE: 2001-08-21
; PRIOR FILING DATE: 2001-08-21
; PRIOR FILING DATE: 2000-08-21
; PRIOR FILING DATE: 2001-05-21
; PRIOR FILING DATE: 2000-05-19
; PRIOR FILING DATE: 2001-03-06
; PRIOR FILING DATE: 2001-03-06
; PRIOR FILING DATE: 2001-03-05
; PRIOR FILING DATE: 2000-04-18
; PRIOR FILING DATE: 2001-04-10
; PRIOR FILING DATE: 2001-04-11
; PRIOR FILING DATE: 2000-04-18
; PRIOR FILING DATE: 2001-09-24
; PRIOR FILING DATE: 2001-09-24
; PRIOR FILING DATE: 2000-09-25
; PRIOR FILING DATE: 2001-11-07
; PRIOR FILING DATE: 2000-11-07
; PRIOR FILING DATE: 2001-03-06
; PRIOR FILING DATE: 2000-03-05
; PRIOR FILING DATE: 2000-03-07
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 2649
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-170-789-39
Query Match 81.6%; Score 2547; DB 6; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 214 ATGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 273
Db 1 ATGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 60
QY 274 GAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATTTGTTGAGCGGTAT 333
Db 61 GAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATTTGTTGAGCGGTAT 120
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGTGCGGATACAGAAATATCATGGCTACATGATG 393
Db 121 TCCTGGAGTCAGCTTAAAGAGCTGTGCGGATACAGAAATATCATGGCTACATGATG 180
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
Db 181 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 240
QY 454 GACAGATCTATTACCTTSCCATGTCTGTGAGAACAGAGAAATACACTGTTTATCT 513
Db 241 GACAGATCTATTACCTTSCCATGTCTGTGAGAACAGAGAAATACACTGTTTATCT 300
QY 514 GAATTTCCCAAACTATCAATAGACGAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Db 301 GAATTTCCCAAACTATCAATAGACGAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG 360
QY 574 GATCTTTTCAGGCAACACTTGGATATGGAATTTCTCGAAGAAAGAACTATTAAGA 633
Db 361 GATCTTTTCAGGCAACACTTGGATATGGAATTTCTCGAAGAAAGAACTATTAAGA 420
QY 634 GAAAGAAACGCAATGGAAACAGTCGGAATTTGCTTTACGATTTACCAAGGAAGTGA 693
Db 421 GAAAGAAACGCAATGGAAACAGTCGGAATTTGCTTTACGATTTACCAAGGAAGTGA 480
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACTGTAAGAGATGGAGGCCCAAGGA 753
Db 481 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACTGTAAGAGATGGAGGCCCAAGGA 540
QY 754 TTTACGCAACAACTTTTAAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813
Db 541 TTTACGCAACAACTTTTAAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 600
QY 814 GATCCAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 873
Db 601 GATCCAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 660
QY 874 TGGATATCTAAACATCGTAACAGAGAAAGAGGAGCTCACTTATGTGCAATGAGCTA 933
Db 661 TGGATATCTAAACATCGTAACAGAGAAAGAGGAGCTCACTTATGTGCAATGAGCTA 720
QY 934 GCCAATGGAAGAGATGCCAGATCGCTGAGTCGCTACCTTTGTTTCTCCAGAGAA 993
Db 721 GCCAATGGAAGAGATGCCAGATCGCTGAGTCGCTACCTTTGTTTCTCCAGAGAA 780
QY 994 TTTGATAGATATTTGGCTATTTGGTGTGTCGCAAAAGCTGAAACAACTCCAGTGGTGT 1053
Db 781 TTTGATAGATATTTGGCTATTTGGTGTGTCGCAAAAGCTGAAACAACTCCAGTGGTGT 840
QY 1054 AAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
Db 841 AAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT 900
QY 1114 ACATCCCTTATGTTGAAACAAAGAGGCGAGATCTTCGTTATCTCTAAACAGGTACA 1173
Db 901 ACATCCCTTATGTTGAAACAAAGAGGCGAGATCTTCGTTATCTCTAAACAGGTACA 960
QY 1174 GCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATTTGATGCTGAAGGAGGATC 1233
```



Db	961	GCATAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGAGAGAGGATC	1020
Qy	1234	ATAGATGTCATAGATAAGAACTAATCAACCTTTTGGAGATTCCTATTTTGAAGGAGTTGAA	1293
Db	1021	ATAGATGTCATAGATAAGAACTAATCAACCTTTTGGAGATTCCTATTTTGAAGGAGTTGAA	1080
Qy	1294	TATATTTGCCAGAGCTGGATGGATCTCTGAGGGAAATATGCTTGGTCCATCTCTACTAGAT	1353
Db	1081	TATATTTGCCAGAGCTGGATGGATCTCTGAGGGAAATATGCTTGGTCCATCTCTACTAGAT	1140
Qy	1354	CGTCTCCAGCTCGCCTACAGATAGTGTGATCTCACTGAAATATTTATCCAGTAGAA	1413
Db	1141	CGTCTCCAGCTCGCCTACAGATAGTGTGATCTCACTGAAATATTTATCCAGTAGAA	1200
Qy	1414	GATGATGTTATCGAAGGCAGAGACTCACTGAGTCAGTCTGATCTCTGAGCCACTA	1473
Db	1201	GATGATGTTATCGAAGGCAGAGACTCACTGAGTCAGTCTGATCTCTGAGCCACTA	1260
Qy	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1533
Db	1261	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1320
Qy	1534	CCCAAAGTCACGAAGAGAAATTCAGTTTATTTTGGCTCTGATGCAAAACAGTTTC	1593
Db	1321	CCCAAAGTCACGAAGAGAAATTCAGTTTATTTTGGCTCTGATGCAAAACAGTTTC	1380
Qy	1594	CGTCATTTATCAAAATTCATCTATTTTAAAGAAAGCAATATAAAGCATCCAGTGT	1653
Db	1381	CGTCATTTATCAAAATTCATCTATTTTAAAGAAAGCAATATAAAGCATCCAGTGT	1440
Qy	1654	GGGCTGCTGCTCCAGTCACTTCAAGTGTCTCTAACAAGAGGAGATAGCAATTTACAGT	1713
Db	1441	GGGCTGCTGCTCCAGTCACTTCAAGTGTCTCTAACAAGAGGAGATAGCAATTTACAGT	1500
Qy	1714	GATGAATGGGAAGTCTTGGCCGCAATGATCTCAATATCCAAATGATGATGAGTCAGAGG	1773
Db	1501	GATGAATGGGAAGTCTTGGCCGCAATGATCTCAATATCCAAATGATGATGAGTCAGAGG	1560
Qy	1774	CTGGTATATTTTGAAGGCCCAAGAATCCCTTTAGAGCATCACTGTGATGATCAGT	1833
Db	1561	CTGGTATATTTTGAAGGCCCAAGAATCCCTTTAGAGCATCACTGTGATGATCAGT	1620
Qy	1834	TACGTAAATCTGGAGAGTGCACAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCG	1893
Db	1621	TACGTAAATCTGGAGAGTGCACAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCG	1680
Qy	1894	ATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953
Db	1681	ATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1740
Qy	1954	GTTGCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAGGAATTT	2013
Db	1741	GTTGCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAGGAATTT	1800
Qy	2014	TGGGCCACCAATTTTGGATTGACAGGCTCTCTCTCTGACTATACCTCTCCAGAAATTTTC	2073
Db	1801	TGGGCCACCAATTTTGGATTGACAGGCTCTCTCTCTGACTATACCTCTCCAGAAATTTTC	1860
Qy	2074	TCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	2133
Db	1861	TCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	1920
Qy	2134	CAGCTGGAAAGAAATCTCTGCTGCTTCAATATGTTGGTGGTCTCAGTGCAGTTG	2193
Db	1921	CAGCTGGAAAGAAATCTCTGCTGCTTCAATATGTTGGTGGTCTCAGTGCAGTTG	1980
Qy	2194	GTTGAATATCGTTTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCCTCTAGGT	2253
Db	1981	GTTGAATATCGTTTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCCTCTAGGT	2040
Qy	2254	TATGTGTTGTAGTGATAGACAAACAGGGGATCCTGTCCAGGAGCTTAAATTTGAAGGC	2313
Db	2314	GCCTTTAATATAAATGGGTCAATAGAAATGACATCAGGTGGAAGGACTCCCAATAT	2373
Qy	2101	GCCTTTAATATAAATGGGTCAATAGAAATGACATCAGGTGGAAGGACTCCCAATAT	2160
Db	2374	CTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCAACGGTGGTCTAT	2433
Qy	2161	CTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCAACGGTGGTCTAT	2220
Db	2434	GGAGGATACCTCTCCCTCATGGCAATTAATCAGAGGTCAGATATCTTCAGGGTTCATAT	2493
Qy	2221	GGAGGATACCTCTCCCTCATGGCAATTAATCAGAGGTCAGATATCTTCAGGGTTCATAT	2280
Db	2494	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTATATG	2553
Qy	2281	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAACGTTATATG	2340
Db	2554	GCTCACCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2613
Qy	2341	GCTCACCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2400
Db	2614	AGTTTCCCTCTGAAACCAATCGTTTACTCTTACTCTTACATGGTTTCTGGATGAGATGTC	2673
Qy	2401	AGTTTCCCTCTGAAACCAATCGTTTACTCTTACTCTTACATGGTTTCTGGATGAGATGTC	2460
Db	2674	CATTTTGCACATACCAAGTATATTTACTAGTGTCTTTTGTAGGGCTGGAAGCCATATGAT	2733
Qy	2461	CATTTTGCACATACCAAGTATATTTACTAGTGTCTTTTGTAGGGCTGGAAGCCATATGAT	2520
Db	2734	TTACAGATCTATCTCAGGAGACACAGCATAGAGTTTCTGAAATCGGAGAACATTAT	2793
Qy	2521	TTACAGATCTATCTCAGGAGACACAGCATAGAGTTTCTGAAATCGGAGAACATTAT	2580
Db	2794	GAATCGATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAAGTATTTGCTCTTAAAA	2853
Qy	2581	GAATCGATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAAGTATTTGCTCTTAAAA	2640
Db	2854	GTGATATAA 2862	
Qy	2641	GTGATATAA 2649	
Db			
RESULT 8			
US-10-054-776-1			
; Sequence 1, Application US/10054776			
; Publication No. US20030165818A1			
; GENERAL INFORMATION:			
; APPLICANT: Mark Robert Edbrooke			
; APPLICANT: Alan Peter Lewis			
; TITLE OF INVENTION: NOVEL PROTEIN			
; FILE REFERENCE: Q61042US			
; CURRENT APPLICATION NUMBER: US/10/054,776			
; CURRENT FILING DATE: 2002-01-23			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 1			
; LENGTH: 2649			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)...(2649)			
US-10-054-776-1			
Query Match			
Best Local Similarity 80.0%; Score 2496; DB 6; Length 2649;			
Matches 2646; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	214	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTCGGACTGT	273
Db	1	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTCGGACTGT	60

Qy	274	GAGGAGAAATATTGAATCA	CAGGATCGGCCCTAAATTTGGAGCCCTTTTAAATGTTGAGCGGTAT	333	
Db	61	GAGGAGAAATATTGAATCA	CAGGATCGGCCCTAAATTTGGAGCCCTTTTAAATGTTGAGCGGTAT	120	
Qy	334	TCCTGGAGTCAGCTTTAAA	AAAGCTGCTTGGCCGATACACAGAAATATCATGGCTACATGATG	393	
Db	121	TCCTGGAGTCAGCTTTAAA	AAAGCTGCTTGGCCGATACACAGAAATATCATGGCTACATGATG	180	
Qy	394	GCTAAGGCACCA	CATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA	453	
Db	181	GCTAAGGCACCA	CATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA	240	
Qy	454	GACAGAATCTATTACCTT	GCCATGCTCGTGTGAGACAGAGAAATACACATGTTTAAATCT	513	
Db	241	GACAGAATCTATTACCTT	GCCATGCTCGTGTGAGACAGAGAAATACACATGTTTAAATCT	300	
Qy	514	GAATTC	CCCAAACTATCAATAGACGACGATCTTAATGCTCTCTGGAAGACCTCTTTTG	573	
Db	301	GAATTC	CCCAAACTATCAATAGACGACGATCTTAATGCTCTCTGGAAGACCTCTTTTG	360	
Qy	574	GATCTTTTTCAGGCA	ACATGACATCTGGATGTAATTTCTCGAGAGGAAGAACTATTAAAG	633	
Db	361	GATCTTTTTCAGGCA	ACATGACATCTGGATGTAATTTCTCGAGAGGAAGAACTATTAAAG	420	
Qy	634	GAAGAAACGCA	TTGGACATCGGTTCCTTACGATTAATCACAGGAAGTGA	693	
Db	421	GAAGAAACGCA	TTGGACATCGGTTCCTTACGATTAATCACAGGAAGTGA	480	
Qy	694	ACATTTCTGTTTC	AAAGCCGATAGTGGAAATTTATCAGTAAAGATGGAAGGCCACAAGGA	753	
Db	481	ACATTTCTGTTTC	AAAGCCGATAGTGGAAATTTATCAGTAAAGATGGAAGGCCACAAGGA	540	
Qy	754	TTTACGCA	CAACCTTTAAGGCCCAATCTAGTGGAAA	CTAGTTGTGCCAACATACGGATG	813
Db	541	TTTACGCA	CAACCTTTAAGGCCCAATCTAGTGGAAA	CTAGTTGTGCCAACATACGGATG	600
Qy	814	GATCCAAA	ATTATGCCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACGATTT	873	
Db	601	GATCCAAA	ATTATGCCCCGCTGATCCAGACTGGATTCCTTTTATACATAGCAACGATTT	660	
Qy	874	TGGATATCT	TAACATCGTAACACAGAGAAGAAAGGAGACTCACTTATGTGCACAAATCAGCTA	933	
Db	661	TGGATATCT	TAACATCGTAACACAGAGAAGAAAGGAGACTCACTTATGTGCACAAATCAGCTA	720	
Qy	934	GCCAACTGA	AGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAA	993	
Db	721	GCCAACTGA	AGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAA	780	
Qy	994	TTTGAT	PAGATATTCTGGCTATTGTTGGTGTGCCAAAGCTGMAACAACCTCCCAAGTGGTGGT	1053	
Db	781	TTTGAT	PAGATATTCTGGCTATTGTTGGTGTGCCAAAGCTGMAACAACCTCCCAAGTGGTGGT	840	
Qy	1054	AAAAATCTT	PAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT	1113	
Db	841	AAAAATCTT	PAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT	900	
Qy	1114	ACATCCCT	TATGTTGAACCAAGAGGCGCATTCATCCGTTATCCTTAAACAGGTACA	1173	
Db	901	ACATCCCT	TATGTTGAACCAAGAGGCGCATTCATCCGTTATCCTTAAACAGGTACA	960	
Qy	1174	GCAATCCT	TAAGTCACTTTTATGATGCTCAGAAATATGATTTGATGCTGAAGGAAGGATC	1233	
Db	961	GCAATCCT	TAAGTCACTTTTATGATGCTCAGAAATATGATTTGATGCTGAAGGAAGGATC	1020	
Qy	1234	ATAGATGCT	ATAGTAAGAACTAATTTCAA	CTTTTGTGAGATTTCTATTTTGAAGGATTTGAA	1293
Db	1021	ATAGATGCT	ATAGTAAGAACTAATTTCAA	CTTTTGTGAGATTTCTATTTTGAAGGATTTGAA	1080
Qy	1294	TATATTG	CCAGCTGGATGGAATCTGAGGGAATAATGTGTGCTCATCTACTAGAT	1353	
Db	1081	TATATTG	CCAGCTGGATGGAATCTGAGGGAATAATGTGTGCTCATCTACTAGAT	1140	
Qy	1354	CGCTCC	CAGACTCGCCCTACAGATAGTGTGTGATCTCACTGAAATTTATTTATCAGTAGAA	1413	

Db	1141	CGCTCCAGACTCGCTGCAGATAGTGTGATCTCACCTGAATATTATTATCCCAAGTAGAA	1200
Qy	1414	GATGATCTTATGGAAGGACAGACATCATTTGATGTCAGTGCCTGATCTGTGACGCCACTA	1473
Db	1201	GATGATGTTATGGAAGGACAGACATCATTTGATGTCAGTGCCTGATCTGTGACGCCACTA	1260
Qy	1474	ATTATCTATGAGAAACAACAGACATCTGGATTAATAATCCATGACATCTTTTCATGTTTT	1533
Db	1261	ATTATCTATGAGAAACAACAGACATCTGGATTAATAATCCATGACATCTTTTCATGTTTT	1320
Qy	1534	CCCAGAGTCCAGAGAGGAAATTCAGTTTATTTTTCCTCTGATGCAAAAACAGGTTTC	1593
Db	1321	CCCAGAGTCCAGAGAGGAAATTCAGTTTATTTTTCCTCTGATGCAAAAACAGGTTTC	1380
Qy	1594	CGTCATTATACAAAATTACATCTATTTTAAAGGAAACAAATATAAAGATCCAGTGGT	1653
Db	1381	CGTCATTATACAAAATTACATCTATTTTAAAGGAAACAAATATAAAGATCCAGTGGT	1444
Qy	1654	GGCTGCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTCACAGT	1713
Db	1441	GGCTGCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTCACAGT	1500
Qy	1714	GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCAAAGTCAAGTTCAGAGG	1773
Db	1501	GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCAAAGTCAAGTTCAGAGG	1560
Qy	1774	CTGATATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTACCTGTCAGT	1833
Db	1561	CTGATATATTTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTACCTGTCAGT	1620
Qy	1834	TAGTAAATCTCGAGAGGTGACAAGGCTGACCGTGCATGACCGTGGCTACTCACTTCCTGTC	1893
Db	1621	TAGTAAATCTCGAGAGGTGACAAGGCTGACCGTGCATGACCGTGGCTACTCACTTCCTGTC	1680
Qy	1894	ATCAGTCAGCATGTGACTCTTTTAAAGTAAAGTATAGTAACGAGAGATCCACACTGT	1953
Db	1681	ATCAGTCAGCATGTGACTCTTTTAAAGTAAAGTATAGTAACGAGAGATCCACACTGT	1740
Qy	1954	GTGTCCTTTTCAAGCATCAAGTCCCTGAAGATGACCCAACTTGCAAAACAAGGAAATTT	2013
Db	1741	GTGTCCTTTTCAAGCATCAAGTCCCTGAAGATGACCCAACTTGCAAAACAAGGAAATTT	1800
Qy	2014	TGGGCCACATTTTGGATTTGACAGGTGCTCTCTCTGACTATACCTCTCCAGAAATTTTC	2073
Db	1801	TGGGCCACATTTTGGATTTGACAGGTGCTCTCTCTGACTATACCTCTCCAGAAATTTTC	1860
Qy	2074	TCCTTTGAAAGTACATCTGATTTTCAATGTATGGATGCTCTCAACAGCCTCATGATCTA	2133
Db	1861	TCCTTTGAAAGTACATCTGATTTTCAATGTATGGATGCTCTCAACAGCCTCATGATCTA	1920
Qy	2134	CAGCTCGAAAGAAATATCTCTACTGTGCTGTTTATATATGTTGGTGCCTCAGGTGCAGTTG	2193
Db	1921	CAGCTCGAAAGAAATATCTCTACTGTGCTGTTTATATATGTTGGTGCCTCAGGTGCAGTTG	1980
Qy	2194	GTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAATACCTCAGCTCTCTAGGT	2253
Db	1981	GTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAATACCTCAGCTCTCTAGGT	2040
Qy	2254	TATGTTGTTAGTATAGACAAACAGGGATCTGTCAACAGGGCTTAAATTTTGAAGGC	2313
Db	2041	TATGTTGTTAGTATAGACAAACAGGGATCTGTCAACAGGGCTTAAATTTTGAAGGC	2100
Qy	2314	GCCTTTAAATATAAATTCGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATAT	2373
Db	2101	GCCTTTAAATATAAATTCGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATAT	2160
Qy	2374	CTAGCTTCTCCATATGATTTTCACTTGTAGTTCGTTGGGCATCCACGGCTGGTCTAT	2433
Db	2161	CTAGCTTCTCCATATGATTTTCACTTGTAGTTCGTTGGGCATCCACGGCTGGTCTAT	2220
Qy	2434	GGAGGATACCTTCCCTGATGGCATTAATGACAGGTCAGATATCTTCAGGGTGTATTT	2493

Db 2221 GAGGATACCTCTCCCTGATGGCAATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTATT 2280  
Qy  
2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553  
Db  
2281 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2340  
Qy  
2554 GGTCAACCTGACAGGATGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613  
Db  
2341 GGTCAACCTGACAGGATGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2400  
Qy  
2614 AAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC 2673  
Db  
2401 AAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC 2460  
Qy  
2674 CATTTTGCACATACCCAGTATATTAATGAGTTTTTTAGTGGGCTGGAAAGCCATATGAT 2733  
Db  
2461 CATTTTGCACATACCCAGTATATTAATGAGTTTTTTAGTGGGCTGGAAAGCCATATGAT 2520  
Qy  
2734 TTACAGATCTATCTCAGAGAGACACAGCATTAAGAGTTCTGAAATCGGGAGAACATTAT 2793  
Db  
2521 TTACAGATCTATCTCAGAGAGACACAGCATTAAGAGTTCTGAAATCTGGAGAACATTAT 2580  
Qy  
2794 GAACCTGATCTTTTGCATCTACCTTCAAGAAACCTTGGATCAGCTATTGCTCTAAAA 2853  
Db  
2581 GAACCTGATCTTTTGCATCTACCTTCAAGAAACCTTGGATCAGCTATTGCTCTAAAA 2640  
Qy  
2854 GTGATATAA 2862  
Db  
2641 GTGATATAA 2649

## RESULT 9

US-09-976-674-2  
; Sequence 2, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 2671  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-2

Query Match 80.0%; Score 2495; DB 3; Length 2671;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 213 CATGGCAGCAGCAATGGAACACAGAACAGCTGGGTGTGAGATATTTGAACTGGGACTG 272  
Db 7 CATGGCAGCAGCAATGGAACACAGAACAGCTGGGTGTGAGATATTTGAACTGGGACTG 66  
Qy 273 TdAGGAGATATGAATCAGAGATCGGCTAAATGGAGCGCTTTTATGTTGAGCGGTA 332  
Db 67 TdAGGAGATATGAATCAGAGATCGGCTAAATGGAGCGCTTTTATGTTGAGCGGTA 126  
Qy 333 TTCTCTGAGTCAAGTAAAGCTGCTTGGCCGATACAGAAATATCATGGCTACATGAT 392  
Db 127 TTCTCTGAGTCAAGTAAAGCTGCTTGGCCGATACAGAAATATCATGGCTACATGAT 186  
Qy 393 GGCTAAGCCACCATGATTTTATGTTGAGAGGAATGATCCAGATGGACCTCATTC 452  
Db 187 GGCTAAGCCACCATGATTTTATGTTGAGAGGAATGATCCAGATGGACCTCATTC 246

Qy 453 AGCAGAACTCTATTCTTCCATGTCTGGTGAGAAACAGAGAAATACACTGTTTTTATTC 512  
Db 247 AGCAGAACTCTATTCTTCCATGTCTGGTGAGAAACAGAGAAATACACTGTTTTTATTC 306  
Qy 513 TGAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAAAGCTCTTTT 572  
Db 307 TGAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAAAGCTCTTTT 366  
Qy 573 GGATCTTTTTCAGGCAACACTGGA CTATGGAATGTATTCTCGAGAGAAAGAACTATTAAAG 632  
Db 367 GGATCTTTTTCAGGCAACACTGGA CTATGGAATGTATTCTCGAGAGAAAGAACTATTAAAG 426  
Qy 633 AGAAGAAACCGATTTGGACAGCTCGGAATGCTCTTACGATATATCACCAAGGAAGTGG 692  
Db 427 AGAAGAAACCGATTTGGACAGCTCGGAATGCTCTTACGATATATCACCAAGGAAGTGG 486  
Qy 693 AACATTTCTCTTTCAAGCCGGTAGTGGAAATTTATCACTGTAAGAGATGAGGGCCACAAGG 752  
Db 487 AACATTTCTCTTTCAAGCCGGTAGTGGAAATTTATCACTGTAAGAGATGAGGGCCACAAGG 546  
Qy 753 ATTTTACGCAACAACTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCTCCAAACATACGGAT 812  
Db 547 ATTTTACGCAACAACTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCTCCAAACATACGGAT 606  
Qy 813 GGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAAGATAT 872  
Db 607 GGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAAGATAT 666  
Qy 873 TTGGATATCTAAACATCGTAAACAGAGAAAGAGAGACTCACTTATGTCACAAATGAGCT 932  
Db 667 TTGGATATCTAAACATCGTAAACAGAGAAAGAGAGACTCACTTATGTCACAAATGAGCT 726  
Qy 933 AGCCAAACATGSAAGAGATGCCAGATCAGCTGAGTGCCTACTCTTGTCTCCAAAGAA 992  
Db 727 AGCCAAACATGSAAGAGATGCCAGATCAGCTGAGTGCCTACTCTTGTCTCCAAAGAA 786  
Qy 993 ATTTGATAGATATTTCTGGCTATTTGGTGTCTCAAAAGCTGAAACAACTCCAGTGGTGG 1052  
Db 787 ATTTGATAGATATTTCTGGCTATTTGGTGTCTCAAAAGCTGAAACAACTCCAGTGGTGG 846  
Qy 1053 TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT 1112  
Db 847 TAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT 906  
Qy 1113 TACATCCCTATGTTGGAAACAGAGAGGAGAGATTCATTCCTGTTATCTTAAACAGGTAC 1172  
Db 907 TACATCCCTATGTTGGAAACAGAGAGGAGAGATTCATTCCTGTTATCTTAAACAGGTAC 966  
Qy 1173 AGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATGCTGTAAGGAAGGAT 1232  
Db 967 AGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATGCTGTAAGGAAGGAT 1026  
Qy 1233 CATAGATGTCATAGATAAGGAACTAAATCAAACCTTTTGAGATTTCTATTGTAAGGAGTTGA 1292  
Db 1027 CATAGATGTCATAGATAAGGAACTAAATCAAACCTTTTGAGATTTCTATTGTAAGGAGTTGA 1086  
Qy 1293 ATATATTGCCAGAGCTGGATGGACTCCTGAGGAAATATGCTTGGTCCATCTACTAGA 1352  
Db 1087 ATATATTGCCAGAGCTGGATGGACTCCTGAGGAAATATGCTTGGTCCATCTACTAGA 1146  
Qy 1353 TCGCTCCAGACTCGCTCTACAGATAGTGTGATCTCACTGAAATTTATTTATCCCAAGTGA 1412  
Db 1147 TCGCTCCAGACTCGCTCTACAGATAGTGTGATCTCACTGAAATTTATTTATCCCAAGTGA 1206  
Qy 1413 AGATGATGTTATGAAAGGAGAGACTCATTTGAGTCAGTCCGCTGATTTCTGTGACCCCACT 1472  
Db 1207 AGATGATGTTATGAAAGGAGAGACTCATTTGAGTCAGTCCGCTGATTTCTGTGACCCCACT 1266  
Qy 1473 AATTATCTATGAACAAACACAGACATCTCGATTAATATCCATGACATCTTTCAATGTTTT 1532  
Db 1267 AATTATCTATGAACAAACACAGACATCTCGATTAATATCCATGACATCTTTCAATGTTTT 1326

QY 1533 TCCCAAAGTCA CGAAGAGAAAATTGAGTTTATTTTGGCTCTGTAATGCAAAACAGGTTT 1592  
Db 1327 TCCCAAAGTCA CGAAGAGAAAATTGAGTTTATTTTGGCTCTGTAATGCAAAACAGGTTT 1386  
QY 1593 CCGTCATTATATACAAAATTACATCTATTTTAAAGGAAAGCAATATAACCATCCAGTGG 1652  
Db 1387 CCGTCATTATATACAAAATTACATCTATTTTAAAGGAAAGCAATATAACCATCCAGTGG 1446  
QY 1653 TGGGCTGCTGCTCCCAAGTCATTTCAAGTGTCTATCAAGGAGGAGATGCAATATACAG 1712  
Db 1447 TGGGCTGCTGCTCCCAAGTCATTTCAAGTGTCTATCAAGGAGGAGATGCAATATACAG 1506  
QY 1713 TGGTGAATGGGAGTTCTTGGCGGCATCGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1772  
Db 1507 TGGTGAATGGGAGTTCTTGGCGGCATCGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1566  
QY 1773 GCTGGTATATTTGAGGACCAAAAGACTCCCTTTAGAGCATACCTGTACGTAGTCAG 1832  
Db 1567 GCTGGTATATTTGAGGACCAAAAGACTCCCTTTAGAGCATACCTGTACGTAGTCAG 1626  
QY 1833 TTACGTAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTG 1892  
Db 1627 TTACGTAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTG 1686  
QY 1893 CATCAGTCAGACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTG 1952  
Db 1687 CATCAGTCAGACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTG 1746  
QY 1953 TGTGTCCCTTTTACAAGCTATCAAGTCTCAAGTCTCGAAGTACCCAACTTGCAAAACAAAGGAAT 2012  
Db 1747 TGTGTCCCTTTTACAAGCTATCAAGTCTCGAAGTACCCAACTTGCAAAACAAAGGAAT 1806  
QY 2013 TTGGGCGCACCATTTTGGATTACAGAGTCTCTCTCGACTATATCTCTCGAGAAATTTT 2072  
Db 1807 TTGGGCGCACCATTTTGGATTACAGAGTCTCTCTCGACTATATCTCTCGAGAAATTTT 1866  
QY 2073 CTCTTTTGAAGTACTACGTGATTTACATCTGTATGGATCTCTACAGGCTCATGATCT 2132  
Db 1867 CTCTTTTGAAGTACTACGTGATTTACATCTGTATGGATCTCTACAGGCTCATGATCT 1926  
QY 2133 ACAGCTCGGAAAGAAATCTCTACTGTGCTGTTCATATATGTTGGTCTCTCAGGTGCAGTT 2192  
Db 1927 ACAGCTCGGAAAGAAATCTCTACTGTGCTGTTCATATATGTTGGTCTCTCAGGTGCAGTT 1986  
QY 2193 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCCCTAGCCCTCTTAGG 2252  
Db 1987 GGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCCCTAGCCCTCTTAGG 2046  
QY 2253 TTATGTGGTTGTAGTATAGACAACAGGGGATCCCTGTACCGAGGGGCTTAAATTTGAAGG 2312  
Db 2047 TTATGTGGTTGTAGTATAGACAACAGGGGATCCCTGTACCGAGGGGCTTAAATTTGAAGG 2106  
QY 2313 CGCCTTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATA 2372  
Db 2107 CGCCTTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATA 2166  
QY 2373 TCTAGCTTCTCGATATGATTTTCATTTGACTTTAGATCGTGTGGGATCCAGGCTGGTCCCTA 2432  
Db 2167 TCTAGCTTCTCGATATGATTTTCATTTGACTTTAGATCGTGTGGGATCCAGGCTGGTCCCTA 2226  
QY 2433 TGAGGATACCTCTCCCTGATGGCATTAAATCCAGGCTCAGATATCTTCAAGGGTGTCTAT 2492  
Db 2227 TGAGGATACCTCTCCCTGATGGCATTAAATCCAGGCTCAGATATCTTCAAGGGTGTCTAT 2286  
QY 2493 TGCCTGGGGCCGAGTCATCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 2552  
Db 2287 TGCCTGGGGCCGAGTCATCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATAT 2346  
QY 2553 GGGTCACCTTGACCAAGATGAACAGGGGCTATATCTTAGGATCTGTGGCCATGCAAGCAGA 2612  
Db 2347 GGGTCACCTTGACCAAGATGAACAGGGGCTATATCTTAGGATCTGTGGCCATGCAAGCAGA 2406  
QY 2613 AAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGATGT 2672

Db 2407 AAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAAATGT 2466  
QY 2673 CCATTTTGCACATACCAAGTATATATCTGAGTTTCTTGTAGGGCTTGGAAAGCCATATGA 2732  
Db 2467 CCATTTTGCACATACCAAGTATATATCTGAGTTTCTTGTAGGGCTTGGAAAGCCATATGA 2526  
QY 2733 TTTACAGATCTATCTCTCAGGAGAGACACAGCATATAAGAGTTCTTGAATCGGGAGAACATTA 2792  
Db 2527 TTTACAGATCTATCTCTCAGGAGAGACACAGCATATAAGAGTTCTTGAATCGGGAGAACATTA 2586  
QY 2793 TGAATGCATCTTTTTCACATACCTTCAAGAAAACCTTGGATCACGTATTTGTGCTCTTAAA 2852  
Db 2587 TGAATGCATCTTTTTCACATACCTTCAAGAAAACCTTGGATCACGTATTTGTGCTCTTAAA 2646  
QY 2853 AGTGATAT 2860  
Db 2647 AGTGATAT 2654  
RESULT 10  
US-10-982-512-2  
; Sequence 2, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Akinesya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; PRIOR APPLICATION NUMBER: 2004-11-05  
; PRIOR FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US/60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 2671  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-982-512-2  
Query Match 80.0%; Score 2495; DB 9; Length 2671;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2645; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 213 CATGGCAGCAGCAATGGAACAGAACAGACAGCTGGGTGTTGAGATATTTGAAACTCGGACTG 272  
Db 7 CATGGCAGCAGCAATGGAACAGAACAGACAGCTGGGTGTTGAGATATTTGAAACTCGGACTG 66  
QY 273 TGAGGAGATATTTGAATCAAGATCGGCTTAAATTTGAGCCCTTTTATCTTGAGCGGTA 332  
Db 67 TGAGGAGATATTTGAATCAAGATCGGCTTAAATTTGAGCCCTTTTATCTTGAGCGGTA 126  
QY 333 TTCTCGAGTCACTTAAAGCTGCTTCCGATACAGAAATATCATGGCTACATGAT 392  
Db 127 TTCTCGAGTCACTTAAAGCTGCTTCCGATACAGAAATATCATGGCTACATGAT 186  
QY 393 GGTAAAGGACCAATATGATTTTCAATGTTTGAAGGAAATGATCCAGATCGACTCATTC 452  
Db 187 GGTAAAGGACCAATATGATTTTCAATGTTTGAAGGAAATGATCCAGATCGACTCATTC 246  
QY 453 AGACAGAAATCTATTACCTTGGCATGTCTGGTGAGAACAGAGAAATACTGTTTATTTC 512  
Db 247 AGACAGAAATCTATTACCTTGGCATGTCTGGTGAGAACAGAGAAATACTGTTTATTTC 306  
QY 513 TGAATTTCCAAAATCTCAATAGAGCAGCAGTCTTAAATGCTCTCTTGGAAAGCTCTTTT 572  
Db 307 TGAATTTCCAAAATCTCAATAGAGCAGCAGTCTTAAATGCTCTCTTGGAAAGCTCTTTT 366

QY 573 GGAATCTTTTTCAGGCAACATCGACTATGGAATGTATTTCTCGAAGAAGAACTATTAAAG 632  
Db 367 GGAATCTTTTTCAGGCAACATCGACTATGGAATGTATTTCTCGAAGAAGAACTATTAAAG 426  
QY 633 AAAAAAGAAACGCAATCGGACAGTCGGAATTTGCTTCTTACGATTTATCACCAGGAGTGG 692  
Db 427 AAAAAAGAAACGCAATCGGACAGTCGGAATTTGCTTCTTACGATTTATCACCAGGAGTGG 486  
QY 693 AACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGGTAAAGATGAGGCGCACAAAGG 752  
Db 487 AACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGGTAAAGATGAGGCGCACAAAGG 546  
QY 753 ATTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCAACATACGGAT 812  
Db 547 ATTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCAACATACGGAT 606  
QY 813 GGATCCAAATATATGCCCGGTGATCCAGACTGGATTTGCTTTTATACATAGCAACAGATAT 872  
Db 607 GGATCCAAATATATGCCCGGTGATCCAGACTGGATTTGCTTTTATACATAGCAACAGATAT 666  
QY 873 TTGGATATCTAAATCTTAACCAAGAGCAAGAGGAGACTCACTTATGTGCAATAGACT 932  
Db 667 TTGGATATCTAAATCTTAACCAAGAGAGGAGACTCACTTATGTGCAATAGACT 726  
QY 933 AdCCAAATATGGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAGAGAA 992  
Db 727 AGCCAAATATGGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAGAGAA 786  
QY 993 ATTTGATAGATATTTCTGGCTATTTGGTGGTCTCCAAAGCTGAAACAACTCCAGTGGTGG 1052  
Db 787 ATTTGATAGATATTTCTGGCTATTTGGTGGTCTCCAAAGCTGAAACAACTCCAGTGGTGG 846  
QY 1053 TAAATTTCTTAGAATTTCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT 1112  
Db 847 TAAATTTCTTAGAATTTCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT 906  
QY 1113 TACATCCCTATGTTGGAACAAAGGAGGAGATTCATTCGGTTATCTCTAAACAGGTAC 1172  
Db 907 TACATCCCTATGTTGGAACAAAGGAGGAGATTCATTCGGTTATCTCTAAACAGGTAC 966  
QY 1173 AGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGAAGAGGAT 1232  
Db 967 AGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGAAGAGGAT 1026  
QY 1233 CATAGATGTCATAGATGAAGAACTAATTCAACTTTTGAATTCATTTTGAAGAGTTGA 1292  
Db 1027 CATAGATGTCATAGATGAAGAACTAATTCAACTTTTGAATTCATTTTGAAGAGTTGA 1086  
QY 1293 ATATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTACTAGTA 1352  
Db 1087 ATATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATGCTTGGTCCATCTACTAGTA 1146  
QY 1353 TGGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGA 1412  
Db 1147 TGGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGA 1206  
QY 1413 AGATGATGTTATGGAAGAGGAGAGACTCAATGAGTCAGTGCCTGATTTCTGTGACGCCACT 1472  
Db 1207 AGATGATGTTATGGAAGAGGAGAGACTCAATGAGTCAGTGCCTGATTTCTGTGACGCCACT 1266  
QY 1473 AATTTATCTATGAAGAAACAAACAGATCTGGATTAATATTCATGATCACTTTTCATGTTTT 1532  
Db 1267 AATTTATCTATGAAGAAACAAACAGATCTGGATTAATATTCATGATCACTTTTCATGTTTT 1326  
QY 1533 TCCCCAAGTCAACAGAGGAATTTGATTTTATTTTGGCTCTGAAATCCAAAAAGGTTTT 1592  
Db 1327 TCCCCAAGTCAACAGAGGAATTTGATTTTATTTTGGCTCTGAAATCCAAAAAGGTTTT 1386  
QY 1593 CCGTCATTTATACAAATACATCTATTTTAAAGGAAGCAAAATATAAACCATCCAGTGG 1652  
Db 1387 CCGTCATTTATACAAATACATCTATTTTAAAGGAAGCAAAATATAAACCATCCAGTGG 1446  
QY 1653 TGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTATCAAAAGAGGAGATAGCAATTAACAG 1712

Db 1447 TGGGCTGCTGCTCCAAAGTGATTTTCAAGTGTCTTATCAAAAGAGGAGATAGCAATTAACAG 1506  
QY 1713 TGETGAATGGGAAGTCTTTCGCGGCAATGATCTTAATATCCAAAGTTGATGAAGTCAGAG 1772  
Db 1507 TGGTGAATGGGAAGTCTTTCGCGGCAATGATCTTAATATCCAAAGTTGATGAAGTCAGAG 1566  
QY 1773 GCTGTATATTTTTCAGGCAACAAAGACTCCCTTTAGAGCATCACCTCTAGCTAGTCAG 1832  
Db 1567 GCTGTATATTTTTCAGGCAACCAAGACTCCCTTTAGAGCATCACCTCTAGCTAGTCAG 1626  
QY 1833 TTAAGTAAATCCTCGAGAGGTGAACAAGCTGA CTGA CCGTGGCTACTCA CATTTCTTGCTG 1892  
Db 1627 TTAAGTAAATCCTCGAGAGGTGAACAAGCTGA CTGA CCGTGGCTACTCA CATTTCTTGCTG 1686  
QY 1893 CATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACATG 1952  
Db 1687 CATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACATG 1746  
QY 1953 TGTGTCCCTTTTACAGCTATCAAGTCTGGAAGATGACCAACTTGCAAAACAAGGAATTT 2012  
Db 1747 TGTGTCCCTTTTACAGCTATCAAGTCTGGAAGATGACCAACTTGCAAAACAAGGAATTT 1806  
QY 2013 TTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGCTACTATCTCTCCAGAAATTTT 2072  
Db 1807 TTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGCTACTATCTCTCCAGAAATTTT 1866  
QY 2073 CTCTTTTGAAGTACTACTGGATTTTACA TTGTATGGGATGCTCTACAAGCCTCATGATCT 2132  
Db 1867 CTCTTTTGAAGTACTACTGGATTTTACA TTGTATGGGATGCTCTACAAGCCTCATGATCT 1926  
QY 2133 ACAGCTGGAAAGAAATATCTTA CTGTGTGTGTTTCAATATATGTTGGTCTCTCAGGTGCTG 2192  
Db 1927 ACAGCTGGAAAGAAATATCTTA CTGTGTGTGTTTCAATATATGTTGGTCTCTCAGGTGCTG 1986  
QY 2193 GGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATATACCTAGACCTCTCTAGG 2252  
Db 1987 GGTGAATATCGA TTTAAAGGAGTCAAGTATTTCCGCTTGAATATACCTAGACCTCTCTAGG 2046  
QY 2253 TTATGTGTTGTAGTATAGACAACAGGAGTCTCTGCTACCGAGGCTTAAATTTGAAG 2312  
Db 2047 TTATGTGTTGTAGTATAGACAACAGGAGTCTCTGCTACCGAGGCTTAAATTTGAAG 2106  
QY 2313 CGCTTTTAAATATATAAATGCTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATA 2372  
Db 2107 CGCTTTTAAATATATAAATGCTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATA 2166  
QY 2373 TCTAGCTTCTCGATATGATTTTCA TTGACTTGTGTTGGGCTCCACCGCTGGTCTCTA 2432  
Db 2167 TCTAGCTTCTCGATATGATTTTCA TTGACTTGTGTTGGGCTCCACCGCTGGTCTCTA 2226  
QY 2433 TGGAGGATACCTCTCTCCCTGATGGCATTTATGACAGAGTTCAGATATCTTCAGGTTGCTAT 2492  
Db 2227 TGGAGGATACCTCTCTCCCTGATGGCATTTATGACAGAGTTCAGATATCTTCAGGTTGCTAT 2286  
QY 2493 TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTGTTAT 2552  
Db 2287 TGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTGTTAT 2346  
QY 2553 GGGTCAACCTGACCAAGATGAACAGGCTATTTACTTTAGGATCTGTGGCCATCGACAGAGA 2612  
Db 2347 GGGTCAACCTGACCAAGATGAACAGGCTATTTACTTTAGGATCTGTGGCCATCGACAGAGA 2406  
QY 2613 AAAAGTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAAATG 2672  
Db 2407 AAAAGTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAAATG 2466  
QY 2673 CCATTTTGCACATACCAAGTATATTA CTGAGTTTTTTTAGTGGGCTGGAAAGCCATATGA 2732  
Db 2467 CCATTTTGCACATACCAAGTATATTA CTGAGTTTTTTTAGTGGGCTGGAAAGCCATATGA 2526  
QY 2733 TTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATTA 2792

Db 2527 TTTACAGATCTATCCTCAGGAGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTA 2586  
Qy 2793 TGAATGCATCTTTTGCACTACCTTCAAGAAACCTTGGATCAGCTATTTGCTCTAAA 2852  
Db 2587 TGAATGCATCTTTTGCACTACCTTCAAGAAACCTTGGATCAGCTATTTGCTCTAAA 2646  
Qy 2853 AGTGATAT 2860  
Db 2647 AGTGATAT 2654

RESULT 11  
US-10-956-157-2177  
; Sequence 2177, Application US/10956157  
; Publication No. US20050118625A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William  
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH  
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES  
; FILE REFERENCE: 031896-043000 (AM 101081)  
; CURRENT APPLICATION NUMBER: US/10/956,157  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 319805  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2177  
; LENGTH: 2830  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-956-157-2177

Query Match 68.4%; Score 2135; DB 9; Length 2830;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2185; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 AAGTCTTAAGCCTCCGAGCCCAAGCGCTGCTACTACCGCGCTGCTTACTAGTCGG 60  
Db 1 AAGTCTTAAGCCTCCGAGCCCAAGCGCTGCTACTACCGCGCTGCTTACTAGTCGG 60  
Qy 61 CGTTCCGCGCTGGGTTGTCAACCGCGCGCGCGAGAGAGCCACTGCAACCGAGCCG 120  
Db 61 CGTTCCGCGCGCTGGGTTGTCAACCGCGCGCGCGAGAGAGCCACTGCAACCGAGCCG 120  
Qy 121 GAGTGGAGCGCGCGAGCATGAAAGCGCGCGAGCGCGCTCCATAGCGCACTCGGAGCG 180  
Db 121 GAGTGGAGCGCGCGAGCATGAAAGCGCGCGAGCGCGCTCCATAGCGCACTCGGAGCG 180  
Qy 181 TCCGGCGGGGCGGGGGGAGGAAATGCAATGCGGCGAGCAATGGAACAGAAACAG 240  
Db 181 TCCGGCGGGGCGGGGGGAGGAAATGCAATGCGGCGAGCAATGGAACAGAAACAG 240  
Qy 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300  
Db 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGATCGG 300  
Qy 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAAGCTTAAAGAGCTGTT 360  
Db 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAAGCTTAAAGAGCTGTT 360  
Qy 361 GCCGATACCAAGAAATATCATGCTACATGATGGCTAAGGACCAACATGATTTTCATGTT 420  
Db 361 GCCGATACCAAGAAATATCATGCTACATGATGGCTAAGGACCAACATGATTTTCATGTT 420  
Qy 421 GTGAAGAGGATGATCCAGATGAGCTCATTCAGACAGAACTTATACCTTGGCATGCT 480  
Db 421 GTGAAGAGGATGATCCAGATGAGCTCATTCAGACAGAACTTATACCTTGGCATGCT 480  
Qy 481 GGTGAGAACAGAGAAATACACTGTTTATTTCTGAANTTCCGAAATTCGAAATAGAGA 540  
Db 481 GGTGAGAACAGAGAAATACACTGTTTATTTCTGAANTTCCGAAATTCGAAATAGAGA 540  
Qy 541 GCAGTCTAATGCTCTCTGGAGAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
Db 541 GCAGTCTAATGCTCTCTGGAGAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600

Db 541 GCAGTCTAATGCTCTCTTTGGAGAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600  
Qy 601 GGAATCTATTTCTCGAGAGAGAGAACTATTTAAGAGAGAGAGAAACGCTATGGACAGTCGA 660  
Db 601 GGAATCTATTTCTCGAGAGAGAGAACTATTTAAGAGAGAGAGAAACGCTATGGACAGTCGA 660  
Qy 661 ATTGCTCTTTACGATTTATCAACGAGAGTGGAACTATTTCTGTTTCAAGCCGGTAGTGGA 720  
Db 661 ATTGCTCTTTACGATTTATCAACGAGAGTGGAACTATTTCTGTTTCAAGCCGGTAGTGGA 720  
Qy 721 ATTTATCAACGTAAGAGATGGAGGCGCAAGGATTTAGCAACAACTTTAAGGCCCAAT 780  
Db 721 ATTTATCAACGTAAGAGATGGAGGCGCAAGGATTTAGCAACAACTTTAAGGCCCAAT 780  
Qy 781 CTAGTGGAAACTAGTTGTCTCCAAACATACGAGATGGATCCAAATTTATGCCCTGCTGATCCA 840  
Db 781 CTAGTGGAAACTAGTTGTCTCCAAACATACGAGATGGATCCAAATTTATGCCCTGCTGATCCA 840  
Qy 841 GACTGGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAACAGAGAA 900  
Db 841 GACTGGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAACAGAGAA 900  
Qy 901 GAAAGGAGAGCTCACTTATGTGCACAAATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960  
Db 901 GAAAGGAGAGCTCACTTATGTGCACAAATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960  
Qy 961 GCTGAGTTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020  
Db 961 GCTGAGTTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020  
Qy 1021 TGTCCTCAAAAGCTGAAACAACTCCAGTGTGTGTAATTTCTAGAAATTTCTATATGAGAA 1080  
Db 1021 TGTCCTCAAAAGCTGAAACAACTCCAGTGTGTGTAATTTCTAGAAATTTCTATATGAGAA 1080  
Qy 1081 AATCATGAATCTGAGTGGGAAATTTTATGTTTATCATCCCTATGTTGGAAACAGAGAGG 1140  
Db 1081 AATCATGAATCTGAGTGGGAAATTTTATGTTTATCATCCCTATGTTGGAAACAGAGAGG 1140  
Qy 1141 GCAGATTCATTCGTTTATCTCTAATAACAGGTAACGAAATCTTAAAGTCACTTTTAAAGTG 1200  
Db 1141 GCAGATTCATTCGTTTATCTCTAATAACAGGTAACGAAATCTTAAAGTCACTTTTAAAGTG 1200  
Qy 1201 TCAGAAATAATGATTTGATGCTGAAGAGGATCATAGATGTCTATAGATAAGAACTAATTT 1260  
Db 1201 TCAGAAATAATGATTTGATGCTGAAGAGGATCATAGATGTCTATAGATAAGAACTAATTT 1260  
Qy 1261 CAACCTTTTGGAGATTTCTATTTGAAGGATTTGAATATTTTGCAGAGCTGGATGGACTCCT 1320  
Db 1261 CAACCTTTTGGAGATTTCTATTTGAAGGATTTGAATATTTTGCAGAGCTGGATGGACTCCT 1320  
Qy 1321 GAGGAAAAATATGCTTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380  
Db 1321 GAGGAAAAATATGCTTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380  
Qy 1381 TTGATCTCAGCTGAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGAGAGACTC 1440  
Db 1381 TTGATCTCAGCTGAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGAGAGACTC 1440  
Qy 1441 ATTGAGTCACTGCTGATTTCTGTGAGCGCACTAAATTTCTATGAAAGAAACAAACAGATC 1500  
Db 1441 ATTGAGTCACTGCTGATTTCTGTGAGCGCACTAAATTTCTATGAAAGAAACAAACAGATC 1500  
Qy 1501 TGGATAAATCATCATGACATCTTTTATGTTTCCCAAGTCCAGAGAGAGAAATTTGAG 1560  
Db 1501 TGGATAAATCATCATGACATCTTTTATGTTTCCCAAGTCCAGAGAGAGAAATTTGAG 1560  
Qy 1561 TTTATTTTGGCTCTGAATGCAACAGGTTTCCGCTCATTTTATACAAATTTACATCTATT 1620  
Db 1561 TTTATTTTGGCTCTGAATGCAACAGGTTTCCGCTCATTTTATACAAATTTACATCTATT 1620  
Qy 1621 TTAAGGAAAGCAATATTAACGATCCAGTGGTGGCTGCTGCTCGAGTGATTTCAAG 1680  
Db 1621 TTAAGGAAAGCAATATTAACGATCCAGTGGTGGCTGCTGCTCGAGTGATTTCAAG 1680





Db 1201 TCAGAAATATGATTGATGCTGAAGGAGCATCATAGATGTCTATAGATAAGAACTAAT 1260  
QY CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGACCTCT 1320  
Db CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGACCTCT 1320  
QY GAGGAAAATATGCTTGGTCCATCTCTACTAGATCCGCTCCAGACGCGCTACATAGTG 1380  
Db GAGGAAAATATGCTTGGTCCATCTCTACTAGATCCGCTCCAGACGCGCTACATAGTG 1380  
QY TTGATCTCACTGAATTTATTTATCCAGTAGAGATGATGATGATGATGATGATGATGATG 1440  
Db TTGATCTCACTGAATTTATTTATCCAGTAGAGATGATGATGATGATGATGATGATGATG 1440  
QY ATTGAGTCAGTGCTGATCTGTGAGCGCCACTAATTTATCTATGAGAAACAAACAGACATC 1500  
Db ATTGAGTCAGTGCTGATCTGTGAGCGCCACTAATTTATCTATGAGAAACAAACAGACATC 1500  
QY TGGATAAATATCCATGACATCTTTCTATGTTTTTCCCAAAGTCCAGAGGAAATTTGAG 1560  
Db TGGATAAATATCCATGACATCTTTCTATGTTTTTCCCAAAGTCCAGAGGAAATTTGAG 1560  
QY TTTATTTTGGCTCTGAATGCAAAACAGGTTTCOGTCAATTTATATCAAAATTTACATCTATT 1620  
Db TTTATTTTGGCTCTGAATGCAAAACAGGTTTCOGTCAATTTATATCAAAATTTACATCTATT 1620  
QY TTAAGGAAAGCAAAATATATAACGATCCAGTGGTGGCTGCTCCCAAGTGATTTCAAG 1680  
Db TTAAGGAAAGCAAAATATATAACGATCCAGTGGTGGCTGCTCCCAAGTGATTTCAAG 1680  
QY TGTCCTATCAAGAGAGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGGCGGCAT 1740  
Db TGTCCTATCAAGAGAGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTTCTTGGCGGCAT 1740  
QY GGAATCTAATATCAAGATTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCAACCAAGAC 1800  
Db GGAATCTAATATCAAGATTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCAACCAAGAC 1800  
QY TCCCTTTAGACATCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
Db TCCCTTTAGACATCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
QY CTGACTGACCGTGGCTACTCACATTTCTGTGCATCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 1920  
Db CTGACTGACCGTGGCTACTCACATTTCTGTGCATCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 1920  
QY AGTAAGTATAGTAACCAAGAAATCCACATCTGTGTCCCTTTTACAGCTATCAAGTCTCT 1980  
Db AGTAAGTATAGTAACCAAGAAATCCACATCTGTGTCCCTTTTACAGCTATCAAGTCTCT 1980  
QY GAAGATGACCAACTTGCMAAACAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040  
Db GAAGATGACCAACTTGCMAAACAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040  
QY CCTCTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100  
Db CCTCTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA 2100  
QY TTGTATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTG 2160  
Db TTGTATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTG 2160  
QY CTGTTCAATATGTTGGTGC 2179  
Db CTGTTCAATATGTTGGTGC 2179

RESULT 13  
US-10-982-512-8  
; Sequence 8, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:

; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; PRIOR FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 4523  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-982-512-8

Query Match 68.2%; Score 2128; DB 9; Length 4523;  
Best Local Similarity 99.9%; Pred No. 0;  
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AAGTGTAAAGCCTCGAGGCCAAGCCGCTGCTACTGCGCCGCTCTTCTTAGTGCCG 60  
Db 1 AAGTGTAAAGCCTCGAGGCCAAGCCGCTGCTACTGCGCCGCTCTTCTTAGTGCCG 60  
QY 61 CGTTCCGCCCTGGTGTGTACCGGCGCGCGCGAGGAGCCACTGCAACCAAGACCG 120  
Db 61 CGTTCCGCCCTGGTGTGTACCGGCGCGCGCGAGGAGCCACTGCAACCAAGACCG 120  
QY 121 GAGTGAGCGCGCGCAGCATGAAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGACCG 180  
Db 121 GAGTGAGCGCGCGCAGCATGAAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGACCG 180  
QY 181 TCCGGCGCGCGCGCGGGGAAAGGAAATGCAACATGCGCAGCAATGGAACAGAACAG 240  
Db 181 TCCGGCGCGCGCGCGGGGAAAGGAAATGCAACATGCGCAGCAATGGAACAGAACAG 240  
QY 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAAATCAGAGGATCGG 300  
Db 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAAATCAGAGGATCGG 300  
QY 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCTTGGAGTCAGCTTAAAGCTGCTT 360  
Db 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCTTGGAGTCAGCTTAAAGCTGCTT 360  
QY 361 GCCGATACAGAAATATCATGGCTACATGATGCTTAAGGCCACCATGATTTCAITGTT 420  
Db 361 GCCGATACAGAAATATCATGGCTACATGATGCTTAAGGCCACCATGATTTCAITGTT 420  
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATACCTGCCATGCT 480  
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATACCTGCCATGCT 480  
QY 481 GGTGAGACAGAGAAATACACTGTTTATTTCTGAATTTCCCAAACTATCAATAGAGA 540  
Db 481 GGTGAGACAGAGAAATACACTGTTTATTTCTGAATTTCCCAAACTATCAATAGAGA 540  
QY 541 GCAGCTTAATGCTCTCTTGGAAAGCCTCTTTGGATCTTTTTCAGGCAACACTGACTAT 600  
Db 541 GCAGCTTAATGCTCTCTTGGAAAGCCTCTTTGGATCTTTTTCAGGCAACACTGACTAT 600  
QY 601 GGAATGATTTCTCGAGAAAGAACTATTAAAGAGAAAGAAACCGATTTGGAACAGTCGGA 660  
Db 601 GGAATGATTTCTCGAGAAAGAACTATTAAAGAGAAAGAAACCGATTTGGAACAGTCGGA 660  
QY 661 ATTGCTTTTACGATTATATCCAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA 720  
Db 661 ATTGCTTTTACGATTATATCCAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA 720  
QY 721 ATTTATCACTAAAGATCGAGGGCCACAAAGGATTTTACGCAACACCTTTTAAGGCCAAT 780



Qy	301	CCTAAATTGGAGCGCTTTTATGTGTAGCGGTATTCTCGAGTCAGCTTTAAAAAGCTGCTT	360
Db	301	CCTAAATTGGAGCGCTTTTATGTGTAGCGGTATTCTCGAGTCAGCTTTAAAAAGCTGCTT	360
Qy	361	GCGGATACCAAGAAAATATCATGGCTACATGATGGCTAAGCCACACATGATTCATGTTT	420
Db	361	GCGGATACCAAGAAAATATCATGGCTACATGATGGCTAAGCCACACATGATTCATGTTT	420
Qy	421	GTGAAGAGGAATGATCCAGATCGAGCCTCATTCAGACAGAACTCTATTACCTTGGCATGTCT	480
Db	421	GTGAAGAGGAATGATCCAGATCGAGCCTCATTCAGACAGAACTCTATTACCTTGGCATGTCT	480
Qy	481	GGTGAGAAACAGAGAAAATACACTGTGTTTATTCGAAATCCGAAAACATCATCAATAGAGCA	540
Db	481	GGTGAGAAACAGAGAAAATACACTGTGTTTATTCGAAATCCGAAAACATCATCAATAGAGCA	540
Qy	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGAACTGATCTTCGAGAGAGAACTATTAGAGAAAAGAAAACGCAATTCGAAACAGTCGGA	660
Db	601	GGAACTGATCTTCGAGAGAGAACTATTAGAGAAAAGAAAACGCAATTCGAAACAGTCGGA	660
Qy	661	ATTGCTCTTACGATATATCAACAAAGGAAGTGGAACTTTCTGTTTCAAGCCCGGTAGTCGA	720
Db	661	ATTGCTCTTACGATATATCAACAAAGGAAGTGGAACTTTCTGTTTCAAGCCCGGTAGTCGA	720
Qy	721	ATTTATCACGTAAAAAGATGGAGGCCCAAGAGATTTCAGCAACAACTTTAAGGCCCAAT	780
Db	721	ATTTATCACGTAAAAAGATGGAGGCCCAAGAGATTTCAGCAACAACTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTGTGTCCTCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAAACTAGTGTGTCCTCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGGATTGCTTTTATACATAGCAACGATATTGGATATCTAACTCGTAACACAGAGAA	900
Db	841	GACTGGATTGCTTTTATACATAGCAACGATATTGGATATCTAACTCGTAACACAGAGAA	900
Qy	901	GAAGGAGACTCACTTTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA	960
Db	901	GAAGGAGACTCACTTTATGTGCAATGAGCTAGCCAACTGGAAGAGATGCCAGATCA	960
Qy	961	GCTGGAGTCGCTACCTTTGTTCTTCCAGAGAAATTTGATAGATATCTTGCTATTTGGTG	1020
Db	961	GCTGGAGTCGCTACCTTTGTTCTTCCAGAGAAATTTGATAGATATCTTGCTATTTGGTG	1020
Qy	1021	TGTCCTAAAAGCTGAACAACCTCCAGTGGTGGTAAATCTTTAGAAATCTATATGAAGAA	1080
Db	1021	TGTCCTAAAAGCTGAACAACCTCCAGTGGTGGTAAATCTTTAGAAATCTATATGAAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGGAAATATTCATGTTACATCCCTATGTTGGAAAACAAAGGAG	1140
Db	1081	AATGATGAATCTGAGGTGGAAATATTCATGTTACATCCCTATGTTGGAAAACAAAGGAG	1140
Qy	1141	GCAGTTCATTCGGTATCTCTAAAACAGGTCAGCAAAATCTTAAAGTCACTTTTAAGATG	1200
Db	1141	GCAGTTCATTCGGTATCTCTAAAACAGGTCAGCAAAATCTTAAAGTCACTTTTAAGATG	1200
Qy	1201	TCAGAAATATGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACATAAT	1260
Db	1201	TCAGAAATATGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAACATAAT	1260
Qy	1261	CAACCTTTGAGATTCCTATTGGAAGGATTGAATATATTGCCAGAGCTGGATGGACTCCT	1320
Db	1261	CAACCTTTGAGATTCCTATTGGAAGGATTGAATATATTGCCAGAGCTGGATGGACTCCT	1320
Qy	1321	GAGGAAAAATGCTTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG	1380
Db	1321	GAGGAAAAATGCTTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG	1380

QY	1381	TTGATCTCACCTGGAATTATTTATCCAGTAGAAGATGATGTATGGAAAGCGCAGAGACTC	1440
DB	1381	TTGATCTCACCTGGAATTATTTATCCAGTAGAAGATGATGTATGGAAAGCGCAGAGACTC	1440
QY	1441	ATTGAGTCAGTGCCTCGATTCTGTGACGCCCACTAATTATCTATGAAGAAACAAACAGACATC	1500
DB	1441	ATTGAGTCAGTGCCTCGATTCTGTGACGCCCACTAATTATCTATGAAGAAACAAACAGACATC	1500
QY	1501	TGGATAAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAAGAGGAAATTGAG	1560
DB	1501	TGGATAAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAAGAGGAAATTGAG	1560
QY	1561	TTTATTTTGGCTCTGGAATCAAAACAGGTTTCGGTCATTATACAAATACATCTATT	1620
DB	1561	TTTATTTTGGCTCTGGAATCAAAACAGGTTTCGGTCATTATACAAATACATCTATT	1620
QY	1621	TTAAGGAAGCAATATNAAGNATCAAGTGTGGCTGCTGCTCCAAGTAGATTCCAAG	1680
DB	1621	TTAAGGAAGCAATATNAAGNATCAAGTGTGGCTGCTGCTCCAAGTAGATTCCAAG	1680
QY	1681	TGTCCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGCGCCGGCAT	1740
DB	1681	TGTCCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGCGCCGGCAT	1740
QY	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAAGCCTGGTATATTTTGAAGGCCACCAAGAC	1800
DB	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAAGCCTGGTATATTTTGAAGGCCACCAAGAC	1800
QY	1801	TCCCCTTTTAGAGCATCACTGTGACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGG	1860
DB	1801	TCCCCTTTTAGAGCATCACTGTGACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGG	1860
QY	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATA	1920
DB	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTATA	1920
QY	1921	AGTAAGTATAGTAACAGAGAATCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCCT	1980
DB	1921	AGTAAGTATAGTAACAGAGAATCCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCCT	1980
QY	1981	GAAGATGACCCAACTTGCAGAAACAAGAAATTTGGGCCACCATTTTGGATTACAGAGT	2040
DB	1981	GAAGATGACCCAACTTGCAGAAACAAGAAATTTGGGCCACCATTTTGGATTACAGAGT	2040
QY	2041	CCCTCTCTGACTATFACTCTCTCGAATAATTTTCTCTTTTGAAGPACTACTGGATTTTACA	2100
DB	2041	CCCTCTCTGACTATFACTCTCTCGAATAATTTTCTCTTTTGAAGPACTACTGGATTTTACA	2100
QY	2101	TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
DB	2101	TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
QY	2161	CTGTTTCATATATGGTGGTC	2179
DB	2161	CTGTTTCATATATGGTGGTC	2179

```

RESULT 15
US-10-982-512-20
; Sequence 20, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP-IV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117

```

; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-20

Query Match 68.2%; Score 2128; DB 9; Length 4676;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAGTCTAAAGCCTCCGAGGCCAAGGCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG 60
DB 1 AAGTCTAAAGCCTCCGAGGCCAAGGCGCTGCTACTGCGCGCTGCTTCTTAGTGCGG 60
QY 61 CQTTCGCCGCTGGGTTGTACACGGCGCGCCGCCGAGGAAGCCACTGCAACACGAGCCG 120
DB 61 CQTTCGCCGCTGGGTTGTACACGGCGCGCCGCCGAGGAAGCCACTGCAACACGAGCCG 120
QY 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGAGGCCGCTCCATAGCGCACGTCGGGACGG 180
DB 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGAGGCCGCTCCATAGCGCACGTCGGGACGG 180
QY 181 TCGGCGCGCGCGCGGGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240
DB 181 TCGGCGCGCGCGCGGGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG 240
QY 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCACAGGATCGG 300
DB 241 CTGGGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCACAGGATCGG 300
QY 301 CTTAAATTTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGTCGTT 360
DB 301 CTTAAATTTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGTCGTT 360
QY 361 GCCGATACAGAAATATCATGGCTACATGATGGCTTAAGGACCAACATGATTTCAATGTTT 420
DB 361 GCCGATACAGAAATATCATGGCTACATGATGGCTTAAGGACCAACATGATTTCAATGTTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTAACCTTGGCCATGTCT 480
DB 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTAACCTTGGCCATGTCT 480
QY 481 GGTGAGAACACAGAGAAATACACTGTTTTTATCTGAAATTTCCCAAACTATCAATAGAGCA 540
DB 481 GGTGAGAACACAGAGAAATACACTGTTTTTATCTGAAATTTCCCAAACTATCAATAGAGCA 540
QY 541 GGAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
DB 541 GGAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
QY 601 GGAATGATTTCTCAGAGAGAGACTATTAAGAGAGAGAAAGCAATTTGGACACTCGGA 660
DB 601 GGAATGATTTCTCAGAGAGAGACTATTAAGAGAGAGAAAGCAATTTGGACACTCGGA 660
QY 661 ATTGCTCTTACGATTTACCAAGGAGTGAACATTTCTGTTCAGAGCGGTAGTGGGA 720
DB 661 ATTGCTCTTACGATTTACCAAGGAGTGAACATTTCTGTTCAGAGCGGTAGTGGGA 720
QY 721 ATTTATACGTTAAAGATGGAGGCCACAGGATTTTACGCAACCAACCTTTTAAGGCCCAAT 780
DB 721 ATTTATACGTTAAAGATGGAGGCCACAGGATTTTACGCAACCAACCTTTTAAGGCCCAAT 780
QY 781 CTAGTGGAAACTAGTGTGTCCTCAACATACGAGATGATCCAAATTTATGCCCGCTGATCCA 840
DB 781 CTAGTGGAAACTAGTGTGTCCTCAACATACGAGATGATCCAAATTTATGCCCGCTGATCCA 840
QY 841 GACTGGATGCTTTTATACATAGCAACGATATTTTGGATATCTAAACATCGTAAACAGAGAA 900
DB 841 GACTGGATGCTTTTATACATAGCAACGATATTTTGGATATCTAAACATCGTAAACAGAGAA 900

QY 901 GAAAGGAGACTCAGCTTATGTGCAACAATGAGCTAGCCAAACATGGAAGAAGATCCAGATCA 960
DB 901 GAAAGGAGACTCAGCTTATGTGCAACAATGAGCTAGCCAAACATGGAAGAAGATCCAGATCA 960
QY 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
DB 961 GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
QY 1021 TGTCCAAAGCTGAAACAACCTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAGAA 1080
DB 1021 TGTCCAAAGCTGAAACAACCTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAGAA 1080
QY 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAACAAGGAGG 1140
DB 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATGTTGGAACAAGGAGG 1140
QY 1141 GCAGATTTCAATTCCTGTTATCTCTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200
DB 1141 GCAGATTTCAATTCCTGTTATCTCTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200
QY 1201 TCAGAAATTAATGATGATGCTGGAAGGAGATCATAGATGTATAGATAAGGAATCAATTT 1260
DB 1201 TCAGAAATTAATGATGATGCTGGAAGGAGATCATAGATGTATAGATAAGGAATCAATTT 1260
QY 1261 CAACTTTTGAGATTTCTATTTGAAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCCT 1320
DB 1261 CAACTTTTGAGATTTCTATTTGAAAGGATTTGAATATATATTTGCCAGAGCTGGATGGACTCCT 1320
QY 1321 GAGGAAATATGCTTGGTCCATCTACTAGATGCGTCCAGAGCTGCCCTACAGATAGTG 1380
DB 1321 GAGGAAATATGCTTGGTCCATCTACTAGATGCGTCCAGAGCTGCCCTACAGATAGTG 1380
QY 1381 TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
DB 1381 TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
QY 1441 ATTGAGTCAGTCCCTGATTTCTGTGAGCCCACTAAATTTCTATGAAGAAACAACAGACATC 1500
DB 1441 ATTGAGTCAGTCCCTGATTTCTGTGAGCCCACTAAATTTCTATGAAGAAACAACAGACATC 1500
QY 1501 TGGATAAATATCCAATGACATCTTTTCATGTTTCCCAAGTCAAGAGAGAAATGAG 1560
DB 1501 TGGATAAATATCCAATGACATCTTTTCATGTTTCCCAAGTCAAGAGAGAAATGAG 1560
QY 1561 TTTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATT 1620
DB 1561 TTTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATT 1620
QY 1621 TTAAGGAAAGCAATATATAACGATCCAGTGGTGGGCTGCCCTGCTCCAGTGATTTCAAG 1680
DB 1621 TTAAGGAAAGCAATATATAACGATCCAGTGGTGGGCTGCCCTGCTCCAGTGATTTCAAG 1680
QY 1681 TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGGGCTGCCCTGCTCCAGTGATTTCAAG 1740
DB 1681 TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGGGCTGCCCTGCTCCAGTGATTTCAAG 1740
QY 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC 1800
DB 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC 1800
QY 1801 TCCCTTTTGAAGCATACCTGTACGTAGTACGTAAATTCCTGGAGAGGTGACAAGG 1860
DB 1801 TCCCTTTTGAAGCATACCTGTACGTAGTACGTAAATTCCTGGAGAGGTGACAAGG 1860
QY 1861 CTGACTGACCGTGGCTACTCAGATTTCTGCTGATCAGTCAGCACTGTGACTTCTTTATA 1920
DB 1861 CTGACTGACCGTGGCTACTCAGATTTCTGCTGATCAGTCAGCACTGTGACTTCTTTATA 1920
QY 1921 AGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGTCCCTTTTCAAGCTATCAAGTCT 1980
DB 1921 AGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGTCCCTTTTCAAGCTATCAAGTCT 1980
QY 1981 GAAGATGACCCCACTTCCAAAACAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGT 2040

Accession	Sequence	Length
Db	1981 GAAGATGACCAACTTCGAAACAAGGAATTTGGGCCACCATTTTGGATTTCAGCAGGT	2040
Qy	2041 CCTCTTCCTGACTATACTCCTCAGAGAAATTTCTCTTTTGAAGATFACCTACTGGATTTTACA	2100
Db	2041 CCTCTTCCTGACTATACTCCTCAGAGAAATTTCTCTTTTGAAGATFACCTACTGGATTTTACA	2100
Qy	2101 TTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAGAAAGAAATATCCTACTGTG	2160
Db	2101 TTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAGAAAGAAATATCCTACTGTG	2160
Qy	2161 CTGTTTCATATATGGTGGTC	2179
Db	2161 CTGTTTCATATATGGTGGTC	2179

Search completed: May 4, 2006, 10:45:48  
Job time : 2199 secs



GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 09:56:31 ; Search time 1177 Seconds  
(without alignments)  
10795.511 Million cell updates/sec

Title: US-10-825-632-2

Perfect score: 3120

Sequence: 1 aagtgttaagcttcgagg.....agaattactataaaaaaaa 3120

Scoring table: **OTACO N22**

Gapop 60.0 , Gapext 60.0

Searched: 9306428 seqs, 2036268586 residues

Word size: 8

Total number of hits satisfying chosen parameters: 18612072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA\_New.\*

1: /SIDSS/ptodata/2/pubpna/US08\_NEW\_PUB.seq1.\*  
2: /SIDSS/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /SIDSS/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /SIDSS/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
5: /SIDSS/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
6: /SIDSS/ptodata/2/pubpna/US09\_NEW\_PUB.seq1.\*  
7: /SIDSS/ptodata/2/pubpna/US09\_NEW\_PUB.seq2.\*  
8: /SIDSS/ptodata/2/pubpna/US09\_NEW\_PUB.seq3.\*  
9: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
10: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq1.\*  
11: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq2.\*  
12: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq3.\*  
13: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq4.\*  
14: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
15: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq1.\*  
16: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq2.\*  
17: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq3.\*  
18: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq4.\*  
19: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq5.\*  
20: /SIDSS/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Length	ID	Description
1	2957	94.8	3143	Sequence 19, Appl
2	2547	81.6	2649	Sequence 21, Appl
3	118	3.8	1346	Sequence 669313,
4	112	3.6	609	Sequence 743558,
5	112	3.6	609	Sequence 743559,
6	73	2.3	612	Sequence 848446,
7	28	0.9	497	Sequence 482467,
8	26	0.8	26	Sequence 543495,
9	25	0.8	25	Sequence 207313,
10	25	0.8	25	Sequence 207314,
11	25	0.8	25	Sequence 207315,
12	25	0.8	25	Sequence 207316,
13	25	0.8	25	Sequence 207317,
14	25	0.8	25	Sequence 207318,

15	25	0.8	25	17	US-11-121-849-207319	Sequence 207319,
16	25	0.8	25	17	US-11-121-849-207320	Sequence 207320,
17	25	0.8	25	17	US-11-121-849-207321	Sequence 207321,
18	25	0.8	25	17	US-11-121-849-207322	Sequence 207322,
19	25	0.8	25	17	US-11-121-849-207323	Sequence 207323,
20	25	0.8	28	10	US-10-310-914A-543509	Sequence 543509,
21	23	0.7	23	10	US-10-310-914A-543503	Sequence 543503,
22	23	0.7	23	10	US-10-310-914A-543506	Sequence 543506,
23	23	0.7	1275	17	US-11-128-061-624	Sequence 624, App
24	23	0.7	1275	17	US-11-128-049-624	Sequence 624, App
25	22	0.7	22	10	US-10-310-914A-543494	Sequence 543494,
26	22	0.7	22	10	US-10-310-914A-543501	Sequence 543501,
27	22	0.7	22	10	US-10-310-914A-543507	Sequence 543507,
28	22	0.7	22	10	US-10-310-914A-543510	Sequence 543510,
29	21	0.7	21	10	US-10-310-914A-543502	Sequence 543502,
30	21	0.7	21	10	US-10-310-914A-543508	Sequence 543508,
31	21	0.7	21	10	US-10-310-914A-543513	Sequence 543513,
32	21	0.7	535	18	US-11-226-869-428	Sequence 428, App
33	21	0.7	611	7	US-09-925-065A-716183	Sequence 716183,
34	21	0.7	611	7	US-09-925-065A-716184	Sequence 716184,
35	21	0.7	976	18	US-11-096-588A-307	Sequence 307, App
36	21	0.7	1981	7	US-09-925-065A-699851	Sequence 699851,
37	21	0.7	1981	7	US-09-925-065A-699852	Sequence 699852,
38	21	0.7	2159	10	US-10-750-185-47863	Sequence 47863, A
39	21	0.7	2159	10	US-10-750-623-47863	Sequence 47863, A
40	21	0.7	2788	9	US-10-505-928-476	Sequence 476, App
41	21	0.7	2814	17	US-11-186-284-54	Sequence 54, Appl
42	21	0.7	2814	18	US-11-245-147-168	Sequence 168, Appl
43	21	0.7	171732	17	US-11-121-086-98	Sequence 98, Appl
44	21	0.7	191150	13	US-10-517-905-19	Sequence 19, Appl
45	20	0.6	20	10	US-10-310-914A-543499	Sequence 543499,

#### ALIGNMENTS

#### RESULT 1

US-11-151-601-19  
; Sequence 19, Application US/11151601  
; Publication No. US20060003413A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Welch, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.  
; FILE REFERENCE: MPI00-054PNCPL0MNDIVM,  
; CURRENT APPLICATION NUMBER: US/11/151,601  
; CURRENT FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: US 10/170,789  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/882,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/212,078  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/934,406  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US01/26052  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,740  
; PRIOR FILING DATE: 2000-08-21

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: BlastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229) ... (2874)
US-11-151-601-19

Query Match          94.8%; Score 2957; DB 17; Length 3143;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 GCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCTGCTTTCGCGC 70
DB 26 GCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCTGCTTTCGCGC 85

QY 71 CTGGGTTGTCAACCGGCGCGCGCGCGCGAGGAGCCACTGCAACCGAGCCGCGAGTGGAGGC 130
DB 86 CTGGGTTGTCAACCGGCGCGCGCGCGAGGAGCCACTGCAACCGAGCCGCGAGTGGAGGC 145

QY 131 GGCCGACATGAGCGCGCGCGCGCGCTCCATAGGCGACCTGCGGACGCTCCGGCGGG 190
DB 146 GGCGACGATGAGCGCGCGCGCGCGCTCCATAGGCGACCTGCGGACGCTCCGGCGGG 205

QY 191 GCCGGGGGAGGAAATGCAACATGCGAGCGAGCAATGGAAACAGACAGCTGGGTGTTG 250
DB 206 GCCGGGGGAGGAAATGCAACATGCGAGCGAGCAATGGAAACAGACAGCTGGGTGTTG 265

QY 251 AGATATTGAAACTGCGGACCTGAGGAGAAATATTGAATCAAGGATCGGCTTAAATTTGG 310
DB 266 AGATATTGAAACTGCGGACCTGAGGAGAAATATTGAATCAAGGATCGGCTTAAATTTGG 325

QY 311 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTTGCCGATACCA 370
DB 326 AGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGTTGCCGATACCA 385

QY 371 GAAATATCATGCTACATGATGGCTTAAGGCAACCATGATTTTATGTTGTAAGAGGA 430
DB 386 GAAATATCATGCTACATGATGGCTTAAGGCAACCATGATTTTATGTTGTAAGAGGA 445

QY 431 ATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTGGTGAGAAC 490
DB 446 ATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCCATGCTGGTGAGAAC 505

QY 491 GAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCAGCAGCTTAA 550
DB 506 GAGAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGAGCAGCAGCTTAA 565

QY 551 TGCTCTTGGAGGCGCTCTTTGGATCTTTTTCAGGGAACACTGGAATATGGAATGTT 610
DB 566 TGCTCTTGGAGGCGCTCTTTGGATCTTTTTCAGGGAACACTGGAATATGGAATGTT 625

QY 611 CTGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCCGAAATGCTTCT 670
DB 626 CTGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCCGAAATGCTTCT 685

QY 671 ACGATTATACCAAGGAAGTGAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACG 730
DB 686 ACGATTATACCAAGGAAGTGAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACG 745

QY 731 TAAAGATGAGGGGCCAAGGATTTACGCAACAACTTTAAGGCCCAACTAGTGGAAA 790
DB 746 TAAAGATGAGGGGCCAAGGATTTACGCAACAACTTTAAGGCCCAACTAGTGGAAA 805

QY 791 CTAGTTGTCCCAACATACGAGTGGATCCAAATTTATGCCCGCTGATCCAGATGGATTG 850
DB 806 CTAGTTGTCCCAACATACGAGTGGATCCAAATTTATGCCCGCTGATCCAGATGGATTG 865

QY 851 CTTTATATACAGCAACGATATTTGGATATCTAACATCGTAAACGAGAGAAAGGAGAC 910
```

```
DB 866 CTTTATATACATAGCAACGATATTTGGATATCTAACATCGTAAACGAGAAAGAGGAGAC 925
QY 911 TCACCTATGTGCACAATAGCTAGCCAAACATGGAAGAAGATCCAGATCAGCTGGAGTCG 970
DB 926 TCACCTATGTGCACAATAGCTAGCCAAACATGGAAGAAGATCCAGATCAGCTGGAGTCG 985
QY 971 CTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAG 1030
DB 986 CTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAG 1045

QY 1031 CTGAAAACAACTCCCAAGTGGTGTAAATTTCTTAGAATTTCTATATCAAGAAATATGATGAAT 1090
DB 1046 CTGAAAACAACTCCCAAGTGGTGTAAATTTCTTAGAATTTCTATATCAAGAAATATGATGAAT 1105

QY 1091 CTGAGGTGGAAATTTATTTCAATGTTTACATCCCTATGTTGGAACAAGAGGAGGAGATTCAT 1150
DB 1106 CTGAGGTGGAAATTTATTTCAATGTTTACATCCCTATGTTGGAACAAGAGGAGGAGATTCAT 1165

QY 1151 TCCGTTATCTTAAACAGGCTACAGCAAACTCCTTAAGTCACTTTTAAGATGTCAGAAATAA 1210
DB 1166 TCCGTTATCTTAAACAGGTAAGCAAACTCCTTAAGTCACTTTTAAGATGTCAGAAATAA 1225

QY 1211 TGATTGATGCTGAAGAGAGATCATAGATGTCATAGATAAGAACTAATTTCAACTTTTGTG 1270
DB 1226 TGATTGATGCTGAAGAGAGATCATAGATGTCATAGATAAGAACTAATTTCAACTTTTGTG 1285

QY 1271 AGATTTCTTTGAGGAGTTGAATATATTTGCGAGAGCTGGATGGAGCTCCTCAGGGGAAAT 1330
DB 1286 AGATTTCTTTGAGGAGTTGAATATATTTGCGAGAGCTGGATGGAGCTCCTCAGGGGAAAT 1345

QY 1331 ATGCTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAC 1390
DB 1346 ATGCTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAC 1405

QY 1391 CTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGAGAGAGACTCATTGAGTCAG 1450
DB 1406 CTGAAATTTATTTATCCAGTAGAAGATGATGTTTATGGAAGGAGAGAGACTCATTGAGTCAG 1465

QY 1451 TGCCGTAATCTGTGAGCGCACTAAATTTATCTATGAAGAAAACAGACATCTGGATAATAA 1510
DB 1466 TGCCGTAATCTGTGAGCGCACTAAATTTATCTATGAAGAAAACAGACATCTGGATAATAA 1525

QY 1511 TCCATGACATCTTTTCTGTTTTCCTCCAAAGTCAGAAAGAGGAAATTTGAGTTTATTTTGTG 1570
DB 1526 TCCATGACATCTTTTCTGTTTTCCTCCAAAGTCAGAAAGAGGAAATTTGAGTTTATTTTGTG 1585

QY 1571 CCTCTGAATGCAAAAACAGGTTTCCGTCATTTTATACAAATTTACATCTATTTTAAAGGAAA 1630
DB 1586 CCTCTGAATGCAAAAACAGGTTTCCGTCATTTTATACAAATTTACATCTATTTTAAAGGAAA 1645

QY 1631 GCAATATATAACGATCCAGTGGTGGCTGCTGCTCAGAGTATTCAGTGTCTCTATCA 1690
DB 1646 GCAATATATAACGATCCAGTGGTGGCTGCTGCTCAGAGTATTCAGTGTCTCTATCA 1705

QY 1691 AAGAGGAGATAGCAATTTACAGTGTGAAATGGGAAGTTCTTTGGCCGGCATGGATCTTAATA 1750
DB 1706 AAGAGGAGATAGCAATTTACAGTGTGAAATGGGAAGTTCTTTGGCCGGCATGGATCTTAATA 1765

QY 1751 TCCAAGTTGATGAAGTCAAGAGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTTTTAG 1810
DB 1766 TCCAAGTTGATGAAGTCAAGAGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTTTTAG 1825

QY 1811 AGCATCACCTGTACGTAGTACGTAAATCTCTGGAAGAGTCAAGGCTGACTGACCC 1870
DB 1826 AGCATCACCTGTACGTAGTACGTAAATCTCTGGAAGAGTCAAGGCTGACTGACCC 1885

QY 1871 GTGGCTACTCACATTTCTTGTGCTCAGTCAGTACGTGACTTCTTTTATAGTAAGTATA 1930
DB 1886 GTGGCTACTCACATTTCTTGTGCTCAGTCAGTACGTGACTTCTTTTATAGTAAGTATA 1945

QY 1931 GTAAACAGAGAAATCCACTGTGTGCTTCTTCAAGCTATCAAGTCTCCTGAGATGACCC 1990
```

Db	1946	GTAAACCAAGAAATCCACACTGTGTGTCCTCTTACAGCTATCAAGTCTCTGAAGATGACC	2005
Qy	1991	CAACTTGCACAAACAAAGGAATTTTGGGCCACACATTTTGGATTCAGAGGTCCTCTTCCTG	2050
Db	2006	CAACTTGCACAAACAAAGGAATTTTGGGCCACACATTTTGGATTCAGAGGTCCTCTTCCTG	2065
Qy	2051	ACTATACTCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACATTCGTATGGGA	2110
Db	2066	ACTATACTCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACATTCGTATGGGA	2125
Qy	2111	TGCTCTACAAGCTCATGATCTACAGCCTGGAAAGAAATATCCCTACTGTCTGCTCATAT	2170
Db	2126	TGCTCTACAAGCTCATGATCTACAGCCTGGAAAGAAATATCCCTACTGTCTGCTCATAT	2185
Qy	2171	ATGGTGGTCTCTCAGGTGCAATTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCT	2230
Db	2186	ATGGTGGTCTCTCAGGTGCAATTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCT	2245
Qy	2231	TGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTATAGACACACAGGGGATCTCTGTC	2290
Db	2246	TGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTATAGACACACAGGGGATCTCTGTC	2305
Qy	2291	ACGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAATAGAAATTCAGC	2350
Db	2306	ACGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAATAGAAATTCAGC	2365
Qy	2351	ATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTATAGATCGTG	2410
Db	2366	ATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTATAGATCGTG	2425
Qy	2411	TGGCATCCACGGCTCGTCTATGGAGTACCTCTCCCTGATGGCATTAATGACAGGCT	2470
Db	2426	TGGCATCCACGGCTCGTCTATGGAGTACCTCTCCCTGATGGCATTAATGACAGGCT	2485
Qy	2471	CAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGATCTTCTATGATA	2530
Db	2486	CAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGATCTTCTATGATA	2545
Qy	2531	CAGGATACCGGAACGTTATATGGGTACCTGACAGAAATGAACAGGGCTTATCTTAG	2590
Db	2546	CAGGATACCGGAACGTTATATGGGTACCTGACAGAAATGAACAGGGCTTATCTTAG	2605
Qy	2591	GATCTGGGCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCTTACTGCTCTTAC	2650
Db	2606	GATCTGGGCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCTTACTGCTCTTAC	2665
Qy	2651	ATGGTTTCTGGATGAGATGTCATTTTGGACATACCATATATCTAGTATTTTCTAG	2710
Db	2666	ATGGTTTCTGGATGAGATGTCATTTTGGACATACCATATATCTAGTATTTTCTAG	2725
Qy	2711	TGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAG	2770
Db	2726	TGAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAG	2785
Qy	2771	TTCTGTAATCGGAGAACATTAATGAATCTGTCATCTTTTGGCATCTTCAAGAAACCTTG	2830
Db	2786	TTCTGTAATCGGAGAACATTAATGAATCTGTCATCTTTTGGCATCTTCAAGAAACCTTG	2845
Qy	2831	GATCAGTATTTGCTCTTAAAGTATATAATTTTGACCTGTGTAGAACTCTCTGCTAT	2890
Db	2846	GATCAGTATTTGCTCTTAAAGTATATAATTTTGACCTGTGTAGAACTCTCTGCTAT	2905
Qy	2891	ACATGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATTCATCATC	2950
Db	2906	ACATGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATTCATCATC	2965
Qy	2951	ACATTTTGAATCGGATGATTAATCTACTCTCTGAAATAAATGTTGTCATCGCAGGG	3010
Db	2966	ACATTTTGAATCGGATGATTAATCTACTCTCTGAAATAAATGTTGTCATCGCAGGG	3025
Qy	3011	GTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATGATA	3070
Db	3026	GTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATGATA	3085
Qy	3071	CATATTTCTGAGAGCCCAAGCAATACCAATAGAAATTTACTAAAAA	3120
Db	3086	CATATTTCTGAGAGCCCAAGCAATACCAATAGAAATTTACTAAAAA	3135
RESULT 2			
US-11-151-601-21			
; Sequence 21, Application US/11151601			
; Publication No. US20060003413A1			
; GENERAL INFORMATION:			
; APPLICANT: Millennium Pharmaceuticals, Inc.			
; APPLICANT: Meyers, Rachel E.			
; APPLICANT: Olandt, Peter J.			
; APPLICANT: Kapeller-Libermann, Rosana			
; APPLICANT: Curtis, Rory A. J.			
; APPLICANT: Williamson, Mark			
; APPLICANT: Weich, Nadine			
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,			
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF			
; FILE REFERENCE: MP100-054PIRCP10MNDIVIM			
; CURRENT APPLICATION NUMBER: US/11/151,601			
; CURRENT FILING DATE: 2005-06-13			
; PRIOR APPLICATION NUMBER: US 10/170,789			
; PRIOR FILING DATE: 2002-06-13			
; PRIOR APPLICATION NUMBER: US 09/797,039			
; PRIOR FILING DATE: 2001-02-28			
; PRIOR APPLICATION NUMBER: PCT/US01/06525			
; PRIOR FILING DATE: 2001-02-28			
; PRIOR APPLICATION NUMBER: US 60/186,061			
; PRIOR FILING DATE: 2000-02-29			
; PRIOR APPLICATION NUMBER: US 09/882,166			
; PRIOR FILING DATE: 2001-06-15			
; PRIOR APPLICATION NUMBER: PCT/US01/19269			
; PRIOR FILING DATE: 2001-06-15			
; PRIOR APPLICATION NUMBER: US 60/212,078			
; PRIOR FILING DATE: 2000-06-15			
; PRIOR APPLICATION NUMBER: US 09/934,406			
; PRIOR FILING DATE: 2001-08-21			
; PRIOR APPLICATION NUMBER: PCT/US01/26052			
; PRIOR FILING DATE: 2001-08-21			
; PRIOR APPLICATION NUMBER: US 60/226,740			
; PRIOR FILING DATE: 2000-08-21			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 45			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 21			
; LENGTH: 2649			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-11-151-601-21			
Query Match 81.6%; Score 2547; DB 17; Length 2649;			
Best Local Similarity 99.9%; Pred. No; 0;			
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	214	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGGAGATATTTGAAACTGCGGACTGT	273
Db	1	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGGAGATATTTGAAACTGCGGACTGT	60
Qy	274	GAGGAGATATTCGAATCACAGGATCGGCTTAAATTTGAGCGCTTTTATGTTGAGCGGTAT	333
Db	61	GAGGAGATATTCGAATCACAGGATCGGCTTAAATTTGAGCGCTTTTATGTTGAGCGGTAT	120
Qy	334	TCCTGGAGTCAGCTTTAAAAAGCTGCTTGCCGATACACAGAAATATCATGGCTACATGATG	393
Db	121	TCCTGGAGTCAGCTTTAAAAAGCTGCTTGCCGATACACAGAAATATCATGGCTACATGATG	180
Qy	394	GCTAAGGCACCAATGATTTTTCATGTTTGTGGAAGAGAAATGATCCAGATGGACCTCATCA	453
Db	181	GCTAAGGCACCAATGATTTTTCATGTTTGTGGAAGAGAAATGATCCAGATGGACCTCATCA	240
Qy	454	GACAGAAATCTATTACCTTGCCATGCTGGTGAGAACAGAGAAATACACTGTTTTTCT	513

Db 241 GACAGAACTCTATTACCTTGGCCATGTCTGGTGAGAACAGAGAAAATACACTGTTTATCT 300  
Qy 514 GAAATTCCTCCAAAACATCAATAGACGACGAGCTTAAATGCTCTCTGGAAGCCTCTTTTG 573  
Db 301 GAAATTCCTCCAAAACATCAATAGACGACGAGCTTAAATGCTCTCTGGAAGCCTCTTTTG 360  
Qy 574 GATCTTTTTCAGGCACACCTGGACTATGGAATGTATCTCGAGAAGAAGAACTATTAA 633  
Db 361 GATCTTTTTCAGGCACACCTGGACTATGGAATGTATCTCGAGAAGAAGAACTATTAA 420  
Qy 634 GAAAGAAAACGCAATGGAACAGCTCGGAATTCCTTACGATTAATCACAAAGGAAGTGA 693  
Db 421 GAAAGAAAACGCAATGGAACAGCTCGGAATTCCTTACGATTAATCACAAAGGAAGTGA 480  
Qy 694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACTGTAAGAAAGATGGAGGCCCAAGGA 753  
Db 481 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACTGTAAGAAAGATGGAGGCCCAAGGA 540  
Qy 754 TTTACGCAACAACTTTAAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813  
Db 541 TTTACGCAACAACTTTAAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 600  
Qy 814 GATCCAAATTTATGCCCCGCTGATCCAGACTGGATTCGTTTATACATAGCAACGATATT 873  
Db 601 GATCCAAATTTATGCCCCGCTGATCCAGACTGGATTCGTTTATACATAGCAACGATATT 660  
Qy 874 TGGATATCTAAACATCGTAACACAGAGAAAGAGAGACTCACTTATGTGCAATAGAGCTA 933  
Db 661 TGGATATCTAAACATCGTAACACAGAGAAAGAGAGACTCACTTATGTGCAATAGAGCTA 720  
Qy 934 GCCAACATGGAAGAAAGATCCAGATCAGCTGGAGTGCCTACTTTGTTCTCCAAAGAA 993  
Db 721 GCCAACATGGAAGAAAGATCCAGATCAGCTGGAGTGCCTACTTTGTTCTCCAAAGAA 780  
Qy 994 TTTGATAGATATTTCTGGCTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1053  
Db 781 TTTGATAGATATTTCTGGCTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840  
Qy 1054 ABAATTCCTAGATATCTATGAGAAATGATGATCTGAGGTGGAATTTATCATGTT 1113  
Db 841 ABAATTCCTAGATATCTATGAGAAATGATGATCTGAGGTGGAATTTATCATGTT 900  
Qy 1114 ACATCCCTATGTTGAAACAAAGAGGGCAGATTCAATCCGTTATCTCTAAACAGGTACA 1173  
Db 901 ACATCCCTATGTTGAAACAAAGAGGGCAGATTCAATCCGTTATCTCTAAACAGGTACA 960  
Qy 1174 GAAATTCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGGAAGGAAGATC 1233  
Db 961 GAAATTCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGGAAGGAAGATC 1020  
Qy 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTTCTATTGGAAGGAGTTGAA 1293  
Db 1021 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTTCTATTGGAAGGAGTTGAA 1080  
Qy 1294 TATATTGCCAGAGCTGGATGGACTCCTGAGGAAAAATATGCTTGGTCCATCTACTAGAT 1353  
Db 1081 TATATTGCCAGAGCTGGATGGACTCCTGAGGAAAAATATGCTTGGTCCATCTACTAGAT 1140  
Qy 1354 CGCTCCAGACTCGCTACAGATGTTGATCTCACCCTGAATTTATTTATCCAGTAGAA 1413  
Db 1141 CGCTCCAGACTCGCTCGAGATAGTTGATCTCACCCTGAATTTATTTATCCAGTAGAA 1200  
Qy 1414 GATGATGTTATGGAAGCAGAGACTCAATGAGTGCCTGATTTCTGTGAGCGCACTA 1473  
Db 1201 GATGATGTTATGGAAGCAGAGACTCAATGAGTGCCTGATTTCTGTGAGCGCACTA 1260  
Qy 1474 ATTATCTATGGAAGAACACAGACATCTGGATTAATATCCATGACATCTTTTCAATGTTTT 1533  
Db 1261 ATTATCTATGGAAGAACACAGACATCTGGATTAATATCCATGACATCTTTTCAATGTTTT 1320  
Qy 1534 CCCCAGTCTCAGAGAGGAATTTAGTTTATTTTGGCTCTGGAATGCAAAAACAGGTTTC 1593

Db 1321 CCCAAAGTCAAGAGAGGAATTTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1380  
Qy 1594 CGTCATTTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAAAACGATCCAGTGGT 1653  
Db 1381 CGTCATTTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAAAACGATCCAGTGGT 1440  
Qy 1654 GGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACCAAGT 1713  
Db 1441 GGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACCAAGT 1500  
Qy 1714 GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGATCAGAAGG 1773  
Db 1501 GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGATCAGAAGG 1560  
Qy 1774 CTGCTATATTTTCAAGAGCACCAGGCTCCCTTTTAGAGCATCACTCTGCTAGTCAAGT 1833  
Db 1561 CTGCTATATTTTCAAGAGCACCAGGCTCCCTTTTAGAGCATCACTCTGCTAGTCAAGT 1620  
Qy 1834 TACGTAATCTCGAGAGGTGACAGGCTGACAGGCTGACTGACCGTGGCTACTACATTTCTTGCTGC 1893  
Db 1621 TACGTAATCTCGAGAGGTGACAGGCTGACAGGCTGACTGACCGTGGCTACTACATTTCTTGCTGC 1680  
Qy 1894 ATCAGTCAGCATCTGACTCTTCTTATAGTATAGTATAGTAAACAGAGATCCACATGT 1953  
Db 1681 ATCAGTCAGCATCTGACTCTTCTTATAGTATAGTATAGTAAACAGAGATCCACATGT 1740  
Qy 1954 GTGCTCCCTTTACAGGCTATCAAGTCTGAGATGACCCAACTTGCAGAAACAAAGGAATTT 2013  
Db 1741 GTGCTCCCTTTACAGGCTATCAAGTCTGAGATGACCCAACTTGCAGAAACAAAGGAATTT 1800  
Qy 2014 TGGGCCACCATTTTGGATTCAGAGGTCTCTTCTCTGACTATATCTCTCAGAAATTTTC 2073  
Db 1801 TGGGCCACCATTTTGGATTCAGAGGTCTCTTCTCTGACTATATCTCTCAGAAATTTTC 1860  
Qy 2074 TCTTTTGAAGTACTACTGATTTTACATTTGTAATGGGATGCTCTCAAGCCCTCATGATCTA 2133  
Db 1861 TCTTTTGAAGTACTACTGATTTTACATTTGTAATGGGATGCTCTCAAGCCCTCATGATCTA 1920  
Qy 2134 CAGCTCTGGAAGAAATATCTACTGCTGTTTATATATGTTGTTGCTCTCAGGTGCAAGTGT 2193  
Db 1921 CAGCTCTGGAAGAAATATCTACTGCTGTTTATATATGTTGTTGCTCTCAGGTGCAAGTGT 1980  
Qy 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGT 2253  
Db 1981 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGT 2040  
Qy 2254 TATGTGTTTGTAGTAGACAAACAGGGGATCCCTGTCAACGAGGCTTAAATTTGAAGGC 2313  
Db 2041 TATGTGTTTGTAGTAGACAAACAGGGGATCCCTGTCAACGAGGCTTAAATTTGAAGGC 2100  
Qy 2314 GCCTTTAAATATAAAATGSGTCAATAGAAATTCACCATCAGGTGGAAGGACTCCATAT 2373  
Db 2101 GCCTTTAAATATAAAATGSGTCAATAGAAATTCACCATCAGGTGGAAGGACTCCATAT 2160  
Qy 2374 CTAGCTTCTCGATATGATTTCAATGACTTAGATCGTGTGGCATCCACGGCTGGTCTAT 2433  
Db 2161 CTAGCTTCTCGATATGATTTCAATGACTTAGATCGTGTGGCATCCACGGCTGGTCTAT 2220  
Qy 2434 GGAGGATCTCTCCCTGATGGCATTAATGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2493  
Db 2221 GGAGGATCTCTCCCTGATGGCATTAATGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2280  
Qy 2494 GCTGGGGCCCCAGTCACTCTCTGATCTCTATGATACAGGATACAGGAAACGTTATATG 2553  
Db 2281 GCTGGGGCCCCAGTCACTCTCTGATCTCTATGATACAGGATACAGGAAACGTTATATG 2340  
Qy 2554 GGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATTCGCAAGCAGAA 2613  
Db 2341 GGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATTCGCAAGCAGAA 2400  
Qy 2614 AAGTTCCTCTGTAACCAAAATCGTTTATCTGCTCTTACATGTTTCTGGATGAGAATGTC 2673  
Db 2401 AAGTTCCTCTGTAACCAAAATCGTTTATCTGCTCTTACATGTTTCTGGATGAGAATGTC 2460



```
; Sequence 818446, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 818446
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-818446

Query Match      2.3%; Score 73; DB 7; Length 612;
Best Local Similarity 100.0%; Pred. No. 1.7e-28;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2667 GAATGTCATTTGGACATACACAGTATATCTAGAGTTTCTAGTGGCGCTGGAAGCC 2726
DB 612 GAATGTCATTTGGACATACACAGTATATCTAGAGTTTCTAGTGGCGCTGGAAGCC 553

QY 2727 ATATGATTTACAG 2739
DB 552 ATATGATTTACAG 540

RESULT 7
US-09-925-065A-482467/c
; Sequence 482467, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR FILING DATE: 2000-11-20
; PRIOR FILING DATE: 2000-11-30
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482467
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-482467

Query Match      0.9%; Score 28; DB 7; Length 497;
Best Local Similarity 100.0%; Pred. No. 0.00093;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; Sequence 543495, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 543495
; LENGTH: 26
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-543495

Query Match      0.8%; Score 26; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 CCAAGGCCGCTGCTACTGCGCGCGCT 46
DB 26 CCAAGGCCGCTGCTACTGCGCGCGCT 1

RESULT 9
US-11-121-849-207313
; Sequence 207313, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
; CURRENT FILING DATE: 2005-05-03
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 207313
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-207313

Query Match      0.8%; Score 25; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2801 ATCTTTGCACTACCTTCAAGAAAA 2825
DB 1 ATCTTTGCACTACCTTCAAGAAAA 25

RESULT 10
US-11-121-849-207314
; Sequence 207314, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: 3684.1
```



; CURRENT APPLICATION NUMBER: US/11/121,849  
; PRIOR FILING DATE: 2005-05-03  
; PRIOR APPLICATION NUMBER: 60/567,949  
; PRIOR FILING DATE: 2004-05-03  
; NUMBER OF SEQ ID NOS: 673904  
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1  
; SEQ ID NO 207314  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-11-121-849-207314

Query Match 0.8%; Score 25; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.033; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

QY 2823 AAACCTTGGATCAGTATTGCTGCT 2847  
|||||  
Db 1 AAACCTTGGATCAGTATTGCTGCT 25

## RESULT 11

US-11-121-849-207315  
; Sequence 207315, Application US/11/121849  
; Publication No. US20050272080A1  
; GENERAL INFORMATION:

; APPLICANT: John Palma

; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S

; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849

; CURRENT FILING DATE: 2005-05-03

; PRIOR APPLICATION NUMBER: 60/567,949

; PRIOR FILING DATE: 2004-05-03

; NUMBER OF SEQ ID NOS: 673904

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 207315

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Homo sapien

US-11-121-849-207315

Query Match 0.8%; Score 25; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.033; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

QY 2835 ACGTATTGCTGCTCTAAAGTGATA 2859  
|||||  
Db 1 ACGTATTGCTGCTCTAAAGTGATA 25

## RESULT 12

US-11-121-849-207316  
; Sequence 207316, Application US/11/121849  
; Publication No. US20050272080A1  
; GENERAL INFORMATION:

; APPLICANT: John Palma

; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S

; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849

; CURRENT FILING DATE: 2005-05-03

; PRIOR APPLICATION NUMBER: 60/567,949

; PRIOR FILING DATE: 2004-05-03

; NUMBER OF SEQ ID NOS: 673904

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 207316

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Homo sapien

US-11-121-849-207316

Query Match 0.8%; Score 25; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.033; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

Best Local Similarity 100.0%; Pred. No. 0.033; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

QY 2865 TTGACCTGTGTAGAACTCTCTGCTA 2889  
|||||  
Db 1 TTGACCTGTGTAGAACTCTCTGCTA 25

## RESULT 13

US-11-121-849-207317  
; Sequence 207317, Application US/11/121849  
; Publication No. US20050272080A1  
; GENERAL INFORMATION:

; APPLICANT: John Palma

; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S

; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849

; CURRENT FILING DATE: 2005-05-03

; PRIOR APPLICATION NUMBER: 60/567,949

; PRIOR FILING DATE: 2004-05-03

; NUMBER OF SEQ ID NOS: 673904

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 207317

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Homo sapien

US-11-121-849-207317

Query Match 0.8%; Score 25; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.033; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

QY 2877 GAACTCTCTGTATACACTGGCTAT 2901  
|||||  
Db 1 GAACTCTCTGTATACACTGGCTAT 25

## RESULT 14

US-11-121-849-207318  
; Sequence 207318, Application US/11/121849  
; Publication No. US20050272080A1  
; GENERAL INFORMATION:

; APPLICANT: John Palma

; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S

; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849

; CURRENT FILING DATE: 2005-05-03

; PRIOR APPLICATION NUMBER: 60/567,949

; PRIOR FILING DATE: 2004-05-03

; NUMBER OF SEQ ID NOS: 673904

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 207318

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Homo sapien

US-11-121-849-207318

Query Match 0.8%; Score 25; DB 17; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.033; Indels 0; Gaps 0;  
Matches 25; Conservative 0; Mismatches 0;

QY 2891 ACACCTGGCTATTAAACCAATGAGG 2915  
|||||  
Db 1 ACACCTGGCTATTAAACCAATGAGG 25

## RESULT 15

US-11-121-849-207319  
; Sequence 207319, Application US/11/121849  
; Publication No. US20050272080A1  
; GENERAL INFORMATION:

; APPLICANT: John Palma

; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S

; FILE REFERENCE: 3684.1

; CURRENT APPLICATION NUMBER: US/11/121,849

; CURRENT FILING DATE: 2005-05-03

; PRIOR APPLICATION NUMBER: 60/567,949

; PRIOR FILING DATE: 2004-05-03

; NUMBER OF SEQ ID NOS: 673904

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 207319

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Homo sapien

US-11-121-849-207319

```
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; PRIOR FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/587,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 207319
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
; US-11-121-849-207319

Query Match      0.8%; Score 25; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2930 GAAACACACAGAAATGATCATCATCAT 2954
         |||||
Db       1 GAAACACACAGAAATGATCATCATCAT 25

Search completed: May 4, 2006, 13:43:40
Job time : 1179 secs
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 03:09:14 ; Search time 90.4 Seconds  
(without alignments)  
3032.880 Million cell updates/sec

Title: US-10-825-632-2

Perfect score: 1030

Sequence: 1 aagtgttaagctccagg.....agaattactaaaaaaa 3120

Scoring tables: 0.0

Xgapop 60.0, Xgapext 60.0

Ygapop 60.0, Ygapext 60.0

Fgapop 60.0, Fgapext 7.0

Delop 60.0, Delext 60.0

Searched: 2443163 seqs, 439378781 residues

Words: 1

Total number of hits satisfying chosen parameters: 4589790

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xip  
-Q=/abss/ABSSWEB\_spool/US10825632/runat\_01052006\_110234\_4411/app\_query.fasta\_1  
-DB=A Geneseq -QFMT=fastan -SUFFIX=oligo.n2p.rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p  
-USER=US10825632 -CGN\_1\_348 -runat\_01052006\_110234\_4411 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database : A\_Geneseq\_21.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	882	85.6	882	4 AAB47187	Aab47187 Human DPP
2	882	85.6	882	5 AAG61591	Abg61591 Human DPP
3	882	85.6	882	5 AAG78415	Aag78415 Amino aci
4	882	85.6	882	5 AAE24170	Aae24170 Human dip
5	882	85.6	882	5 AAU74749	Aau74749 Human pro
6	882	85.6	882	5 ADI17086	Adi17086 Human NOV
7	882	85.6	882	6 ABU07720	Abu07720 Human ser
8	706	68.5	746	8 ADI16334	Adi16334 Human pro
9	706	68.5	831	6 ABU92026	Abu92026 Human pro

10	700	68.0	824	6 ABU92030	Abu92030 Human pro
11	658	63.9	782	5 ABB97361	Abb97361 Novel hum
12	655	63.6	658	5 ABG61600	Abg61600 Human DPR
13	655	63.6	661	5 ABG61596	Abg61596 Human DPR
14	655	63.6	690	5 ABG61594	Abg61594 Human DPR
15	607	58.9	613	5 ABG61601	Abg61601 Human DPR
16	563	54.7	587	5 ADR41398	Adr41398 Human CD-
17	476	46.2	724	5 ABB97362	Abb97362 Novel hum
18	406	39.4	632	4 AAB93565	Aab93565 Human pro
19	341	33.1	465	4 AAB47189	Aab47189 Human DPP
20	339	32.9	358	5 ABG61597	Abg61597 Human DPR
21	275	26.7	738	8 ADT04044	Adt04044 Human pro
22	274	26.6	493	7 ADE78977	Ade78977 Human pro
23	262	25.4	360	4 AAB47190	Aab47190 Human DPP
24	250	24.3	250	5 ABB99949	Abb99949 Dipeptidyl
25	238	23.1	241	5 ABG61595	Abg61595 Human DPR
26	194	18.8	516	6 ABU92029	Abu92029 Human pro
27	194	18.8	580	5 AAE14337	Aae14337 Human pro
28	186	18.1	883	5 ADI17085	Adi17085 Murine NO
29	183	17.8	194	5 ABG61599	Abg61599 Human DPR
30	176	17.1	310	4 AAB47188	Aab47188 Human DPP
31	176	17.1	310	5 ABB08994	Abb08994 Human dip
32	176	17.1	310	7 ADD27858	Add27858 Human dip
33	151	14.7	193	5 ABB89739	Abb89739 Human pol
34	136	13.2	136	4 AAB74673	Aab74673 Human pro
35	86	8.3	108	5 ABG61598	Abg61598 Human DPR
36	79	7.7	129	4 AAM94906	Aam94906 Human rep
37	19	1.8	19	5 ABG61613	Abg61613 Human DPR
38	17	1.7	18	4 AAB47191	Aab47191 DPP8 pept
39	15	1.5	15	5 ABB99950	Abb99950 Dipeptidyl
40	15	1.5	16	4 AAB47192	Aab47192 DPP8 pept
41	13	1.3	432	6 ABP97157	Abp97157 Human dip
42	13	1.3	460	4 AAM38740	Aam38740 Human pol
43	13	1.3	512	4 AAM40526	Aam40526 Human pol
44	13	1.3	518	3 AAY90299	Aay90299 Human rep
45	13	1.3	518	5 ABG61603	Abg61603 Human DPR

#### ALIGNMENTS

RESULT 1

AAB47187

ID AAB47187 standard; protein; 882 AA.

AC AAB47187;

XX

XX 29-JUN-2001 (first entry)

DT

XX

DE Human DPP8.

XX

KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;  
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;  
KW growth hormone deficiency; glucose level; mucosal regeneration;  
KW non-insulin dependent diabetes mellitus; glucose intolerance;  
KW immunosuppression.

XX Homo sapiens.

OS

XX

XX Key Location/Qualifiers

FT Active-site 739

FT /note= "Forms part of Ser-Asp-His catalytic triad"

FT Active-site 817

FT /note= "Forms part of Ser-Asp-His catalytic triad"

FT Active-site 849

FT /note= "Forms part of Ser-Asp-His catalytic triad"

XX WO200119866-A1.

XX

XX 22-MAR-2001.

XX

PF 11-SEP-2000; 2000WO-AU001085.

XX

PR 10-SEP-1999; 99AU-00002762.

PR 18-FEB-2000; 2000AU-00005709.  
XX (UNSY ) UNIV SYDNEY.  
XX PI Abbott CA, Gorell MD;  
XX WPI; 2001-281520/29.  
XX N-PSDB; AAC85694.  
XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving  
PT substrates, identifying inhibitors of DPP8 catalytic activity which have  
PT therapeutic uses, and for detecting activated T cells.  
XX Claim 1; Fig 2; 78pp; English.  
XX This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has  
CC substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.  
CC Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase,  
CC because it is capable of hydrolysing the peptide bond C-terminal to Pro  
CC in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is  
CC useful for cleaving a substrate, and for detecting an activated T cell  
CC which involves measuring the level of DPP8 gene expression in a T cell.  
CC The level of DPP8 expression is detected by detecting the amount of DPP8  
CC RNA in the cell. It is also useful for identifying a molecule capable of  
CC inhibiting DPP8 catalytic activity of the substrate by DPP8. Molecules identified as  
CC inhibiting DPP8 catalytic activity may be useful for treating diabetes,  
CC growth hormone deficiency, lowering glucose levels in non-insulin  
CC dependent diabetes mellitus and other disorders involving glucose  
CC intolerance, enhancing mucosal regeneration and as immunosuppressants  
XX Sequence 882 AA;  
SQ  
Alignment Scores:  
Pred. No.: 0 Length: 882  
Score: 882.00 Matches: 882  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 85.6% Indels: 0  
DB: 4 Gaps: 0  
US-10-825-632-2 (1-3120) x AAB47187 (1-882)  
QY 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGTGAGATATTGAAACATGCGGACTGT 273  
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGAGATATTGAATCACAGGATCGGCCTAAATGGAGCCCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAenileGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTGCAGATACAGAAATATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGATGATCCAGATGGACCTCATTTCA 453  
Db 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 CACAGATCTTACCTTGCATGCTGTGTGAGACAGAGAAATACACTGTTTATCT 513  
Db 81 AspArgileTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer 100  
QY 514 GAAATCCCAAACTATCAATAGACGACGAGCTTAATGCTCTCTTGAAGCCCTCTTTTG 573  
Db 101 GluileProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCGAGAGAGAACTATTAAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140  
QY 634 GAAAGAAAACCGCATGGAAACAGTCGGAATTCCTTACGATTTATCACCAGGAAGTGA 693  
Db 141 GluArgLysArgileGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCCCGTAGTGAAATTTATCAGTTAAAAAGATGGAGGCCCAACGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTACGCAACCAACCTTTAAGGCCCAATCTAGTGGAACACTAGTTGTCCCAACATACGGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCCAAAATATTATGCCCGCTGATCCAGACTCGGATGCTTTTATATCATAGCAACCATATT 873  
Db 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTAACATCGTAACACAGAGAAGAGAGACTCATTATGTGCACAAATCAGCTA 933  
Db 221 TrpIleSerAenileValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeu 240  
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTGGCTTACCTTGTCTCCCAAGAACAA 993  
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260  
QY 994 TTTGATAGATATTCTGGCTATTGGTGTCTGCCAAAGCTGGAACACTCCCAAGTGGTGGT 1053  
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATCTTAGAATCTATATGAAGAAGATGATCTGAGCTGGAATATTTCATGCTT 1113  
Db 281 LysileLeuArgileLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300  
QY 1114 ACATCCCTATGTTGGAACAAGAGGGCAGATTCTATCCGTTATCTCTAAACAGGTACA 1173  
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAAACTCTAAAGTCACTTTTAAAGATGTTCAGAAATAATGATGTGCTGAAGGAGGATC 1233  
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATCTCATAGATAAGCAACTAATTCAACCTTTTGGAGATCTATTGAAAGGAGTGA 1293  
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
QY 1294 TATATTCACAGAGCTGGATGGACTCCTGAGGAAAATATGCTGTGCTCATCTACTAGAT 1353  
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380  
QY 1354 CGCTCCACAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATATTATCCCAAGTAGAA 1413  
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATCATCTTATGAAAGGAGGAGACTCATTTGAGTCAGTGCCTGATTCGTGACGCCACTA 1473  
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATCAAGAAACACACACATCTGGATAAATATCCATGACATCTTTCATGTTT 1533  
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
QY 1534 CCCCAAGATCACAAAGGAATTTGATTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1593  
Db 441 ProGlnSerHisGluGluGluIlePheIlePheAlaSerGluCysLysThrGlyPhe 460  
QY 1594 CGTCATTTATACAAAATTCATCTATTTTTAAAGAAACAAATATAAACGATCCAGTGGT 1653  
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
QY 1654 GGGCTGCTCTCCCAAGTCAATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAGT 1713  
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500  
QY 1714 GGTGAATGGGAAGCTTTTGGCCGGCATCGATCTAATATCAAGTTGATGAAGTCAGAGG 1773  
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520

1774 CTGGTATATTTTGAAGCACAAAGACTCCCTTTAGAGCATCACTCTAGCTAGTCAGT 1833  
1775 |||||  
521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540  
1834 TACCTAAATCTCGAGAGGTGACAAGCGTGAAGCCGTTGGCTACTACATCTCTGCTGC 1893  
541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560  
1894 ATCAGTCAGCACTGTGCTCTTTTATTAAGTAACTAGTAACTAGTAACTAGTAACTAGT 1953  
561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
1954 GTGTCCTCTTACAACTATCAAGTCTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2013  
581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
2014 TGGGCCCAACCTTTGGATTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2073  
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620  
2074 TCTTTTCAAGTACTACTGCTATTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTA 2133  
621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
2134 CAGCTCGAAGAAATATCTACTGTGCTGTTTATATATGTTGGTCTCTCAGGTGCAAGT 2193  
641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660  
2194 GTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT 2253  
661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680  
2254 TATCTGTTGTAGTAGTACACACAGGGATCTGTCCAGGGGCTTAAATTTGAAGGC 2313  
681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700  
2314 GCCTTAAATATAAATGGTCAAAATAGAAATGACGATCAGGTGGAAGGACTCCAATAT 2373  
701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720  
2374 CTAGCTTCTGATATGATTTTCATTTAGATCTGTGGGATCCAGGCTGCTCTAT 2433  
721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740  
2434 GAGGATACCTCTCCCTGATGGCAATTAATGACGAGTCAAGATCTTCAAGGTTGCTATT 2493  
741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760  
2494 GTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACCGAAGCTTATATG 2553  
761 AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMet 780  
2554 GTTACCTTACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCTATCGGCAATCGAAGCA 2613  
781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800  
2614 AAGTTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGAATGTC 2673  
801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820  
2674 CATTTTGACATACACATATATTACTAGTGTGTTTATGTCAGGCTGGAAGCATATGAT 2733  
821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840  
2734 TTACAGATCTATCTCAGGAGAGACACAGCATAAAGTTCCTGATCGGAGCATATAT 2793  
841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860  
2794 GAATGATCTTTTGGCACTACCTTCAAGAAACCTTTGGATCAGTATTCGCTCTCAAAA 2853  
861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880  
2854 GTGATA 2859

Db 881 ValIle 882  
RESULT 2  
ABG61591  
ID ABG61591 standard; protein; 882 AA.  
XX  
AC ABG61591;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE Human DPPIV related serine protease DPPP-1.  
XX  
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;  
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;  
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;  
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;  
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;  
KW dyskinesia; reproductive disorder; inflammatory disorder;  
KW metabolic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200231134-A2.  
XX  
PD 18-APR-2002.  
XX  
PF 12-OCT-2001; 2001WO-US031874.  
XX  
PR 12-OCT-2000; 2000US-0240117P.  
XX  
PA (FERR ) FERRING BV.  
XX  
PI Qi S, Akinsanya KO, Riviere PJ, Junien J;  
XX  
DR WPI; 2002-444178/47.  
XX  
DR N-PSDB; ABK83322.  
XX  
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding  
PT the proteins, useful for treating e.g: fungal, bacterial, protozoan and  
PT viral infections, cancers, allergies, neurological disorders, or pain.  
XX  
PS Claim 17; Fig 1; 113pp; English.  
XX  
CC The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins  
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic  
CC acids encoding them are useful for treating infections such as fungal  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591  
CC -ABG61612 represent human DPPP proteins  
XX  
SQ Sequence 882 AA;  
Alignment Scores:  
Pred. No.: 0 Length: 882  
Score: 882.00 Matches: 882  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 85.6% Indels: 0  
DB: Gaps: 5  
US-10-825-632-2 (1-3120) x ABG61591 (1-882)  
QY 214 ATGCAGCAGCAATGGAACAGACAGCAGCTGGGTGTTGAGATATTTGAACTCGGACTGT 273

Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGAGAAATATTGAATCACAGGATCGGCCTTAAATTCGAGCCCTTTTATGTGAGCGGTAT 333  
Db 21 GluGluAenIleGluSerGlnAspArgProLyLeuGluProPheTyValGluArgTyr 40  
QY 334 TCCTGGAGTCAGCTATAAAGCTGCTTGGCGATACACGAAATATCATCGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLySylsLeuAlaAapThrArgLyTyRhiSgLyTyRwMet 60  
QY 394 GCTAAGGCACACATGATTTTCATGCTTGTGAGAGGAATGATCCAGATGGACCTCATCA 453  
Db 61 AlalysAlaProHisAspPheMetPheValLyArgAsnAspProAspGlyProHisSer 80  
QY 454 GACAGATCTATTACCTTGCATGCTGTGTGAGAACAGAGAAAATACACTGTTTATCT 513  
Db 81 AspArgIleTyTyLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyRser 100  
QY 514 GAATTTCCAAACATCAATAGACGACGAGCTTTAATGCTCTCTTGGAAAGCCCTTTTG 573  
Db 101 GluIleProLySylsThrIleAenArgAlaIaValLeuMetLeuSerTrpLySerProLeu 120  
QY 574 GATCTTTTCAGCAACACTGGACTATGGAATGTATCTTCGAGAAGAACTATTAAAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
QY 634 GAAAGAAAACGCAATGGAAACAGTCGGAATTTGCTTACGATTAATCAACAAAGGAAGTGA 693  
Db 141 GluArgLySargIleGlyThrValGlyIleAlaSerTyrAspTyrRhiSgLyTySerGly 160  
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATACAGTAAAGATGGAGGCCACAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyRhiSvalLyAspGlyGlyProGlnGly 180  
QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTACTTCTCCCAACATACCGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAenLeuValGluThrSerCysProAenIleArgMet 200  
QY 814 GATCCAAAATATTAGCCCCCTGATCCAGACTGGATGCTTTTATACATACACAGATATT 873  
Db 201 AspProLyLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAenAspIle 220  
QY 874 TGATATCTAACATCTACACAGAGAAAGAGACTCACTTATGTGCGCAATGAGCTA 933  
Db 221 TrpIleSerAenIleValThrArgGluGluArgArgLeuThrTyRValHisAenGluLeu 240  
QY 934 GCTAACATGGAGAGATGCCAGATCAGCTGGAGTGCCTACCTTGTCTTCCAAAGAGAA 993  
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260  
QY 994 TTTGATAGATATTCTGGCTATTGGTGTCCAAAGCTGAAACAACTCCCACTCCAGTGTGT 1053  
Db 261 PheAspArgTyRserGlyTyRTrpTrpCysProLySylsAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATCTTAGAAATCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113  
Db 281 LysIleLeuArgIleLeuTyRLeuGluAenAspGluSerGluValGluIleIleHisVal 300  
QY 1114 ACATCCCTTATGTTGAAACAGGAGGCGAGATTCATTCGGTTATCTTAAACACAGGTACA 1173  
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyRProLySylsThrGlyThr 320  
QY 1174 GCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATTCATCTCAAGAGAGGATC 1233  
Db 321 AlaAsnProLySylsValThrPheLySylsMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGATAAGAACTAATTCACACTTTTGAGATTTCTATTTTCAAGAGTTTGA 1293  
Db 341 IleAspValIleAspLySylsGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
QY 1294 TATATTGCCAGAGCTGGATGACTCCCTGAGGGAATATGCTTGGTCCATCTACTAGAT 1353  
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLySylsTyRAlaTrpSerIleLeuLeuAsp 380

QY 1354 CGCTCCCAAGACTCCCTTACAGATAGTGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413  
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATGATGTTATGGAAAGCGCAGAGACTCATTTGAGTCAGTGCCTGTGATCTGTGACCGCACTA 1473  
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATGAAGAAAACACACACATCTCGATAAATATCCATGACATCTTTTCATGTTTTT 1533  
Db 421 IleIleTyRLeuGluThrThrAspIleTyrPheAenIleHisAspIlePheHisValPhe 440  
QY 1534 CCCCAAGTCAACAAGAGGAAATTTGAGTTTATTTTTCCTCTGATGATGCAAAACAGGTTTC 1593  
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLyThrGlyPhe 460  
QY 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACCAATATAACAGATCCAGTGGT 1653  
Db 461 ArgHisLeuTyRlySylsIleThrSerIleLeuLySylsSerLyThrLySylsSerGly 480  
QY 1654 GGGCTGCTCTCCAAAGTGATTTTCAAGTGTCTTCAAGAGGAGATAGCAATTTACAGT 1713  
Db 481 GlyLeuProAlaProSerAspPheLySylsCysProIleLySylsGluIleAlaIleThrSer 500  
QY 1714 GGTCAATGGGAAGTCTTCCCGCGCATCGATCTAATATCAAGTTGATGAAGTCAAGAGG 1773  
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAenIleGlnValAspGluValArg 520  
QY 1774 CTGGTATTTTGAAGGCCAACAGACTCCCTTTTAGACATCACTGTAGTAGTCAGT 1833  
Db 521 LeuValTyRLeuGlyThrLySylsSerProLeuGluHisHisLeuTyRValValSer 540  
QY 1834 TAGTAAATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTCG 1893  
Db 541 TyRValAsnProGlyGluValThrArgLeuThrAspArgGlyTyRSerHisSerCysCys 560  
QY 1894 ATCAGTCAGACTGTGACTCTTTTATAAGTATAGTAAACCAAGAGAAATCCACACTGT 1953  
Db 561 IleSerGlnHisCysAspPhePheIleSerLySylsTyRSerAenGlnLySylsAsnProHisCys 580  
QY 1954 GTCTCCTTTTACAAGCTATCAAGTCTCGAAGTCAACCAACTTCAAAACAAAGAGAAATTT 2013  
Db 581 ValSerLeuTyRlySylsLeuSerSerProGluLeuAspProThrCysLyThrLySylsGluPhe 600  
QY 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073  
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyRThrProProGluIlePhe 620  
QY 2074 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA 2133  
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyRLeuIleTyRlySylsProHisAspLeu 640  
QY 2134 CAGCCTGGAAAGAAATATCTCTGCTGTTTCATATATGTTGGTCTCAGGTGCAATGTC 2193  
Db 641 GlnProGlyLySylsTyRProThrValLeuPheIleTyRlyGlyProGlnValGlnLeu 660  
QY 2194 GTCAATATACGGTTTAAAGGATCAAGTNTTCCGCTTGAATACCTTAGCTCTCTAGGT 2253  
Db 661 ValAsnAsnArgPheLySylsGlyValTyRtyRPhenArgLeuAenThrLeuAlaSerLeuGly 680  
QY 2254 TATGCTGTAGTACATAGACACAGGGATCTCTGTCACCGAGGCTTAAATTTGAAGGC 2313  
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLySylsPheGluGly 700  
QY 2314 GCCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCAGTGTGAAGACTCCAATAT 2373  
Db 701 AlaPheLySylsTyRLeuMetGlyGlnIleAspAspGlnValGluGlyLeuGlnTyR 720  
QY 2374 CTAGCTTCTCGATATGATTTTCATGACTTAGTTCGTGTGGCATCCACGGCTGGTCTTAT 2433  
Db 721 LeuAlaSerArgTyRAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyR 740



QY 2434 GAGGATACCTCTCCCTGATGGCATTAATGACAGAGTGCAGATATCTTCAGGGTTCGTATT 2493  
 DB 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnAAspPheArgValAlaIle 760  
 QY 2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGACGTATATG 2553  
 DB 761 AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMet 780  
 QY 2554 GGTCACTCCCTACAGATGACAGGCTATTACTTACGATCTGTGGCCATGACGACGAA 2613  
 DB 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800  
 QY 2614 AAGTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAAATGTC 2673  
 DB 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820  
 QY 2674 CATTTTCACATACCACTATATTACTGAGTCTTTTATGAGCGGTGGAACCCATATGAT 2733  
 DB 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840  
 QY 2734 TTACAGATCTATCTCAGGAGACACAGCATAAAGATGTTCTGAAATCGGAGAACATTAT 2793  
 DB 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860  
 QY 2794 GAAGTGCATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGATTGCTGCTCAAAA 2853  
 DB 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880  
 QY 2854 GTGATA 2859  
 DB 881 ValIle 882

RESULT 3  
 AAG78415  
 ID AAG78415 standard; protein; 882 AA.  
 XX AC AAG78415;  
 XX DT 12-APR-2002 (first entry)  
 XX DE Amino acid sequence of 21953 human prollyl oligopeptidase.  
 XX KW 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;  
 XX KW cardiovascular disease; autoimmune disease; atopic allergy;  
 XX KW neuronal disorder; vascular disorder; prostate disorder; cystostatic;  
 XX KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;  
 XX KW diabetes mellitus; arthritis; multiple sclerosis; asthma;  
 XX KW Grave's disease; neuronal disorder; demyelinating disease.  
 XX OS Homo sapiens.  
 XX PN WO200179473-A2.  
 XX PD 25-OCT-2001.  
 XX PF 11-APR-2001; 2001WO-US040483.  
 XX PR 18-APR-2000; 2000US-0197508P.  
 XX PA (MILL-) MILLENNIUM PHARM INC.  
 XX PI Meyers RA, Williamson M;  
 XX DR WPI; 2002-034353/04.  
 XX DR N-PSDB; AAH99934.  
 XX PT New polypeptides 21953, member of human prollyl oligopeptidase family,  
 XX PT useful as diagnostic targets and therapeutic agents for controlling  
 XX PT cancer, lymphoma and leukemia.  
 XX PS Claim 1; Page 102-103; 12ipp; English.  
 XX CC This invention relates to an isolated 21953 human prollyl oligopeptidase.

CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,  
 CC antithyroid, dermatological, antipsoriatic, antiasthmatic,  
 CC ophthalmological, antiinflammatory, motropic, antiparkinsonian,  
 CC anticonvulsant, gynaecological, vasotropic, antianginal, cardiant,  
 CC antiatherosclerotic, anorectic and metabolic in its action. Uses include  
 CC gene therapy, expression or activity of 21953 protein modulator, it is  
 CC useful for identifying a compound which binds to it and can be used in  
 CC preventing, treating or detecting a cellular proliferative or  
 CC differentiative disorder. The 21953 molecules can act as novel diagnostic  
 CC targets and therapeutic agents for controlling disorders associated with  
 CC the aberrant activity or degradation of peptide hormones e.g., disorders  
 CC associated with cell differentiation and proliferation such as cancer,  
 CC immune function, reproductive, neurological and cardiovascular function.  
 CC The 21953 molecules are thus useful for treating and preventing cellular  
 CC proliferative and differentiative disorders, haematopoietic neoplastic  
 CC disorders, immune disorders such as autoimmune diseases, diabetes  
 CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,  
 CC neuronal disorders, demyelinating diseases, vascular disorders and  
 CC metabolism or pain disorders. This sequence represents the amino acid  
 CC sequence of 21953 human prollyl oligopeptidase  
 XX  
 SQ Sequence 882 AA;  
 Alignment Scores:  
 Pred. No.: 0 Length: 882  
 Score: 882.00 Matches: 882  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 85.6% Indels: 0  
 DB: 5 Gaps: 0  
 US-10-825-632-2 (1-3120) x AAG78415 (1-882)  
 QY 214 ATGGCAGCAGCAATGGAACACAGACAGCTGGTGTGAGATATTGAAACTGCGACTGT 273  
 DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
 QY 274 GAGGAGAAATATTGAATCACAGAGTACGCGCTTAAATTTGGAGCCCTTTTATTTGAGCGTAT 333  
 DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
 QY 334 TCCTGGAGTCAGCTTAAAGCTGTTGCGATACACAGAAATATCATCGGTACATGATG 393  
 DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
 QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATCA 453  
 DB 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
 QY 454 GACAGAACTATTACCTTCCATGCTGCTGGTGAACAGAGAAATACACTGTTTATTCT 513  
 DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
 QY 514 GAAATTTCCAAAACATCATAGCAGCAGGAGCTTAAATGCTCTCTTGGAGCCCTCTTTTG 573  
 DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
 QY 574 GATCTTTTTCAGGCACACACTGGACTATGGAATGTATTCTCGAGAGAAAGAACTATTAGA 633  
 DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
 QY 634 GAAAGAAAACGCAATGGAACAGTCGGAATTCGCTTCTTACGATATATCACCAGGAAGTGA 693  
 DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
 QY 694 ACATTTCTGTTTCAAGCCGGTAGTAGGAAATATATCAGTAAAGATGGAGGGCCACAGA 753  
 DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
 QY 754 TTTACGCAACACCTTTTAAGCCCAATCTAGTGGAACTAGTTGTCACCAACATCGGATG 813  
 DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

QY 814 GATCCAAATTTATGCCCGCTGATCCAGACTGCTTTCATATACATAGCAACGATATT 873  
DB |||||  
QY 201 AEPProLysLeuCyProAlaAEPProAEPTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTTAACATCGTAACACAGAGAGAAAGAGACTCACATCTGTGCACATGAGCTA 933  
DB |||||  
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrValHisAsnIleLeu 240  
QY 934 GCCACATGGGAAGAGATGCCAGATCAGCTGGAAGTCCCTACCTTTGTTCTCCACAGAGAA 993  
DB |||||  
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGluValAlaThrPheValLeuGluGlu 260  
QY 994 TTTGATAGATATCTCGCTATTTGGTGTGCCAAAAGCTGAAACAACTCCACAGTGGTGT 1053  
DB |||||  
QY 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrProSerGlyGly 280  
QY 1054 AAAATCTTAGAATCTATATGAAGAAATGATGAATCTGAGCTGGAATATTATCATGTT 1113  
DB |||||  
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300  
QY 1114 ACATCCCTATGTTGGAACAAGAGGCGAGATTCATTCCTGTTATCTCTAAACAGGTACA 1173  
DB |||||  
QY 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAATCCTAAAGTCACTTTTAAGATGTCAGAAATATGATTGATGCTGGAAGGAGATC 1233  
DB |||||  
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCTATTTGAAGGAGTTGAA 1293  
DB |||||  
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
QY 1294 TATATTGCCAGAGCTGGAGCTCCTGAGGGAATATGCTGCTCCATCTCTACTAGAT 1353  
DB |||||  
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380  
QY 1354 CGCTCCAGAGCTCGCTACAGACTGCTTCATCTCACCTGAATATTATTTATCCAGTAGAA 1413  
DB |||||  
QY 391 ArgSerGlnThrArgLeuGluIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATCATGTTATGGAAGGCGAGACTCATTTGAGTGCAGTGCCTGATCTCTGAGCCCACTA 1473  
DB |||||  
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATGAAGAAACAAGACATCTCGATAAATATCCATGACATCTTTTCATGTTTTT 1533  
DB |||||  
QY 421 IleIleTyrGluGluThrAspIleIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
QY 1534 CCCCAAGTCCAGAGAGGAATTTGAGTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593  
DB |||||  
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
QY 1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT 1653  
DB |||||  
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
QY 1654 GGGCTGCTCTCCAAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713  
DB |||||  
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500  
QY 1714 GGTGAATGGGAAGTCTTTCGGCGGATGGATCTATATCCCAAGTTTGATGAGTCAGACAGG 1773  
DB |||||  
QY 501 GlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGluValArgArg 520  
QY 1774 CTGGTATATTTTCAAGGCCCAAGACTCCCTTTAGAGCATCACCTGTACCTAGTACGT 1833  
DB |||||  
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540  
QY 1834 TAGCTAAATCTGGAGAGGTGACAGGCTGACTGACCTGGCTACTCACATTTCTGCTGC 1893  
DB |||||  
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560  
QY 1894 ATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAAACGAGAAATCCACACTGT 1953

DB |||||  
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
QY 1954 GTGTCCCTTTTCAAGAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGGAATTT 2013  
DB |||||  
QY 581 ValSerLeuTyrLysLeuSerProGluLysAspAspProThrCysLysThrLysGluPhe 600  
QY 2014 TGGGCCCACTTTTGGATTCCAGAGGTCTCTTCTCTGACTATATCTCTCCGAGAAATTTTC 2073  
DB |||||  
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620  
QY 2074 TCTTTTGAAGTACTACTCGATTACATTTGATGTGATGCTCTACAAGCCTCATGATCTA 2133  
DB |||||  
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
QY 2134 CAGCTCGGAAGAAATATCTCTGCTGCTTTCATATATGTTGGTCTCTCAGGTGAGTGTG 2193  
DB |||||  
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu 660  
QY 2194 GTGAATATCGTGTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT 2253  
DB |||||  
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680  
QY 2254 TATCTGTTCTGATGATAGACACAGGGATCTCTGTCACCGAGGCTTTAAATTTGAAGGC 2313  
DB |||||  
QY 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700  
QY 2314 GCCTTTAAATATAAATGGTCAAAATAGAAATGACGATCAGGTGGAAGGATCCAAATAT 2373  
DB |||||  
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720  
QY 2374 CTAGCTCTCGATATGATTTTCATTTAGATCTGATGCTGTTGGGATCCACGGTGTCTAT 2433  
DB |||||  
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyr 740  
QY 2434 GGAGATACCTCTCCCTGATGATTAATGACAGGTGATGATATCTTCAGGGTTGCTATT 2493  
DB |||||  
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760  
QY 2494 GCTGGGCCCTGCTCTGATCTCTGATCTCTATATGATACAGGATACACGGAACGTTATATG 2553  
DB |||||  
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780  
QY 2554 GGTACCCCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613  
DB |||||  
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800  
QY 2614 AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAATGTC 2673  
DB |||||  
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820  
QY 2674 CATTTTGCATACACAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGAT 2733  
DB |||||  
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840  
QY 2734 TTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTTCCTGAATCGGAGACATATAT 2793  
DB |||||  
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860  
QY 2794 GAACTGATCTTTTGGCTACCTTCAAGAAACCTTGGATCCATGCTATTTGCTGCTTAAAA 2853  
DB |||||  
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluLeuGlySerArgIleAlaLeuLeuLys 880  
QY 2854 GTGATA 2859  
DB |||||  
QY 881 ValIle 882  
RESULT 4  
AAE24170 standard; protein; 882 AA.  
ID AAE24170  
XX AAE24170;  
AC AAE24170;  
XX

DT 23-SEP-2002 (first entry)  
DE Human dipeptidyl peptidase 8 (DPP8) protein.  
XX  
KW Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;  
KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;  
KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;  
KW antitviral; enzyme.  
XX  
OS Homo sapiens.  
XX  
PN WO200234900-A1.  
XX  
PD 02-MAY-2002.  
XX  
PF 29-OCT-2001; 2001WO-AU001388.  
XX  
PR 27-OCT-2000; 2000AU-00001078.  
XX  
PA (UNSY ) UNIV SYDNEY.  
XX  
PI Abbott CA, Gorrell MD;  
XX  
DR WPI: 2002-454646/48.  
DR N-PSDB; AD38956.  
XX  
XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors  
PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,  
PT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV  
PT infection.  
XX  
XX Example; Fig 1; 91pp; English.  
PS  
XX  
CC The present invention relates to dipeptidyl peptidase (DPP) proteins and  
CC polynucleotides encoding such proteins. The DPP peptides are useful for  
CC screening inhibitors of DPP catalytic activity. The inhibitors are useful  
CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft  
CC rejection and HIV (human immuno deficiency virus) infection. The present  
CC sequence is human DPP8 protein  
XX  
SQ Sequence 882 AA;  
  
Alignment Scored:  
Pred. No.: 0 Length: 882  
Score: 882.00 Matches: 882  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 85.6% Indels: 0  
DB: 5 Gaps: 0  
  
US-10-825-632-2.(1-3120) x AAE241170 (1-882)  
  
QY 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGGAACATGCGGACTGT 273  
DB 1 MetAlaAlaAlaMetGluThrGluGluLeuGlyValGluIlePheGluThrAlaAspCys 20  
  
QY 274 GAGGAGATATTGAATCACAGGATCGGCCTAAATTTGGAGCTTTTATGTTGAGCGGTAT 333  
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
  
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTCCCGATACCAGAAAAATATCATGGCTACATGATG 393  
DB 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
  
QY 394 GTTAAGGCCACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453  
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
  
QY 454 GACAGAACTTATTACCTGTCATGCTGGTGAGACAGAGAAAAATACACTGTTTATCTCT 513  
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
  
QY 514 GAAATCCCAAAATCATATAGACGAGAGCTCTTAATGCTCTCTTGGAGCGCTTTTGG 573

Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTTCAGGCAACACACTGGACTATGGAATCTATTCTCGAGAGAGAACTATTAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgIleGluGluLeuLeuArg 140  
QY 634 GAAAGAAAACCATTTGGAACAGTCGGAATTCCTTCTTACGATTATCACCAGGAAGTGA 693  
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGCTAAAAAGATGGAGGGCCACAAGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTACGCAACCAACCTTTAAGCCCAATCTAGTCGAAACCTAGTTGTCCTCCCAACATACGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGATTCCTTTTATACATAGCAACCATATT 873  
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTTAACATCGTAACACAGAGAGAAAGAGACTCATTATGTGCACAAATGAGCTA 933  
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240  
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTGTCTCCCAAGAGAA 993  
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGlnGlu 260  
QY 994 TTTGATAGATATTCTGGCTATTGGTGTCTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 1053  
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATTTCTAGAATTTCTATATGAAGAAATATGAAATCTGAGTGGAAATATTATCATGT 1113  
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300  
QY 1114 ACATCCCTATGTTGGAAACAGAGGGCGAGATTCTTCGTTATCTCTTAAACACAGTACA 1173  
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAAAATCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATTCATGCTCAAGAGAGATC 1233  
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGATAAGGAACTAAATCAACCTTTTGGAGATTCTATTTCGAAGAGTTGAA 1293  
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
QY 1294 TATATTGTCAGAGCTGGATGACTCTCTGAGGAAAAATATGCTTGTTGCTCACTACTAGAT 1353  
Db 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380  
QY 1354 CGCTCCACAGATCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413  
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATGATGTTATGGAAGGCGAGAGACTCATTTGAGTCTGCTGATCTGTGACGCCACTA 1473  
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATGAAGAAACACAGACATCTGATAAATATCATGATCATCTTTCATGTTTTT 1533  
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
QY 1534 CCCCAAGTCAAGAGAGGAAATTTGAGTTTATTTTTCCTCTGAAATGCAAAAACAGTTTC 1593  
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
QY 1594 CQTCAATTTATCAAAATTTACATCTATTATTTAAAGGAAGCAAAATATAAACAGTCCAGTGGT 1653

Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
Qy 1654 GGGCTGCTCCAGTGATTTCAAGTGTCTTCAAGAGAGAGAGATTCAGTATTCAGT 1713  
Db 481 GlyLeuProAlaProSerAspPheLysPheProIleLysGluGluIleAlaIleThrSer 500  
Qy 1714 GGTGAATGGGAGTCTTGGCCGCGCATGGATCTAATATCAAGTGTGATGATGATGAGG 1773  
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520  
Qy 1774 CTGGTATATTTGAAGGACCAAGAAGCTCCCTTTAGAGCATCACTGTAGTCAGT 1833  
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540  
Qy 1834 TACGTAATCTCGAGAGGTGACAAAGCTGACCTGACCGTGGCTACTCACATTTCTGTGC 1893  
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgLysTyrSerHisSerCysCys 560  
Qy 1894 ATCAGTCAGCCTGTGCTCTTTTATAGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1953  
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
Qy 1954 GTGTCCTTTTACAAGCTATCAAGTCTGGAAGTACCACTGTCGAAACAAAGGAATTT 2013  
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
Qy 2014 TGGGACACATTTTGGATTACAGAGGTCCTCTTCTGACTATCTACTCTCCAGAAATTTTC 2073  
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620  
Qy 2074 TCTTTTGAAGTACTACTGGAATTTACATGTATGGATGCTCTACAAGCCTCATGATCTA 2133  
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
Qy 2134 CAGCTCGAAGAAATATCTACTGCTGCTTTCATATATGGTGTCTCAGTCAGTCAGTTG 2193  
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu 660  
Qy 2194 GTGATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGT 2253  
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680  
Qy 2254 TATGTGTTGTAGTGATAGACAAAGGAGATCTGTCCAGGAGGCTTAAATTTGAAGGC 2313  
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700  
Qy 2314 GCCTTTAAATATAAATGGTCAAAATGAAATGAGATTCAGTGGAGGACTCCAATAT 2373  
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720  
Qy 2374 CTAGCTCTCGATATGATTTTCAATTGATCTGATGCTGGGATCCACGGCTGCTCAT 2433  
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740  
Qy 2434 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTCAAGATATCTTCAAGGTTGCTATT 2493  
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760  
Qy 2494 GCTGGGCCCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553  
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780  
Qy 2554 GGTCACTCACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCATCAAGCAGAGAA 2613  
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuLysSerValAlaMetGlnAlaGlu 800  
Qy 2614 AAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGC 2673  
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820  
Qy 2674 CATTTTGCATACCACTATTTACTAGTTTCTTGTAGGCTGAGGCTGGAAGCCATATGAT 2733  
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValAlaGalaGlyLysProTyrAsp 840

Qy 2734 TTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCTCGAATCGGAGAACTAT 2793  
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860  
Qy 2794 GAATCGATCTTTTGGCATCTACCTTCAAGAAAACTTTGGATCAGTATTTGCTCTTAAAA 2853  
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880  
Qy 2854 GTGATA 2859  
Db 881 Vallie 882  
RESULT 5  
ID AAU74749 standard; protein; 882 AA.  
XX AAU74749;  
AC AAU74749;  
DT 09-APR-2002 (first entry)  
XX Human protease PRTS-9 protein sequence.  
DE Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
KW cell proliferative disorder; developmental disorder; epilepsy;  
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
KW reproductive disorder; endometriosis.  
XX Homo sapiens.  
OS WO200198468-A2.  
PN 27-DEC-2001.  
DD 13-JUN-2001; 2001WO-US019178.  
PF 16-JUN-2000; 2000US-0212336P.  
PR 22-JUN-2000; 2000US-0213955P.  
PR 29-JUN-2000; 2000US-0215396P.  
PR 07-JUL-2000; 2000US-0216821P.  
PR 14-JUL-2000; 2000US-0218946P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;  
PI Delegeane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;  
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
PI Kallick DA;  
XX WPI; 2002-090437/12.  
DR N-PSDB; ABK12892.  
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in  
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative  
PT (e.g. cancer) disorders.  
XX Claim 1; Page 140-142; 177pp; English.  
XX The present invention relates to twenty one new human proteases, referred  
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the  
CC invention are useful in the diagnosis, treatment and prevention of  
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's  
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial  
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency  
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.  
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,  
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's  
CC disease and reproductive e.g. infertility and endometriosis disorders.  
CC Numerous other examples of each disorder are given in the specification.

CC The present protein sequence represents the human protease PRNS-9 protein		
XX of the invention		
SQ Sequence 882 AA;		
Alignment Scores:		
Pred. No.:	0	Length: 882
Score:	882.00	Matches: 882
Percent Similarity:	100.0%	Conservative: 0
Best Local Similarity:	100.0%	Mismatches: 0
Query Match:	85.6%	Indels: 0
DB:	5	Gaps: 0
US-10-825-632-2 (1-3120) x AAU74749 (1-882)		
QY	214	ATGGCAGCAGCAATGGAACAGACAGCTGGGTTCGATATTTGAAATCGCGACTGT 273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY	274	GAGGAGAATATTGAATCAGAGATCGGCTAAATTTGAGCCCTTTTATGTTGAGCGGTAT 333
DB	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY	334	TCTCGAGTCAGCTTAAAGCTGCTGCCATACCAAGAAATATCATGCTCATGATG 393
DB	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY	394	GCTAAGCCACCATGATTTCAATGTTGTGAAGAGGAATGATCCAGATGACCTCATTC 453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY	454	GACAGAACTTATACCTTCCCATGCTGCTGAGACAGAGAAATACACTGTTTATCT 513
DB	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY	514	GAAATCCCCAAACTATCAATAGCAGCAGCTTAAATGCTCTCTTGGAGCCCTTTTG 573
DB	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTTCGAGAAGAAGAACTATTAGA 633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY	634	GAAAGAAAACGCAATGGAACAGTCGGAATTCCTTACGATATTCACCAAGAGTGA 693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY	694	ACATTTCTGTTTCAGCCGCTAGTGAATTTATCAGTAAAGATGGAGGGCCACAAGGA 753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY	754	TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG 813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY	814	GATCCAAATTTATCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY	874	TGGATATCAATCGTAACAGAGAGAAAGAGACTCATTATGTGTCACAAATGAGCTA 933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY	934	GCACACATGGAAGAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCAGAAGAA 993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluLeu 260
QY	994	TTTGTAGATATTTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCGAGTGGT 1053
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY	1054	AAAAATCTTAGAATTTCTATATGAAGAAATATGAAATCTGAGGTGGAATTTATTCATGTT 1113

DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY	1114	ACATCCCCTATGTTGGAACAACAGAGAGGCGAGATTTCATTCGTTATCTCTAAACACAGGTACA 1173
DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY	1174	GCAAAATCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGATCTCTGAAGAAGGATC 1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY	1234	ATAGATGTCATAGATAAGAACTAATTCACCTTTTTCAGATTCTTATTTGAAGAGATTGAA 1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY	1294	TATATTGCCAGAGCTGGATGGACTCTCGAGGAAATATGCTTGGTCCATCTTACTAGAT 1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY	1354	CGCTCCCGACTCGCTACAGATAGTGTGATCTCACTCGAATTTATTTATCCAGTAGAA 1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY	1414	GATGATGTTTATGAAAGGCAGAGACTCATTGAGTCAGTGCCTGATTTCTGTACGCCACTA 1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY	1474	ATTATCTATCAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY	1534	CCCCAAAGTCACAGAGGAAATTCAGTTATTTTTCCTCTGATCTGAAATCCAAACAGGTTTC 1593
DB	441	ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY	1594	CGTCATTTATACAAATATACATCTATTATAAGGAAGCAATAATAAACCATCCAGTGGT 1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY	1654	GGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGAGATAGCAATTTACAGT 1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY	1714	GGTGAATGGCAAGTCTTTCGCGGATGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
QY	1774	CTGTATATTTTGAAGGCCAACAGACTCCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
DB	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY	1834	TACCTAATCCTCGAGAGGTGACAAGGCTGACTGACCTGGCTACTCACATTTCTTGCTGC 1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY	1894	ATCAGTCAGACACTGTGACTTTCTTTTATAAGTAAGTAGTAACCAAGAGAATCCACACTGT 1953
DB	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY	1954	GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAGGAATTT 2013
DB	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY	2014	TGGGCGCCACATTTTGGATTTCAGAGGTCCTCTCTCTGACTACTACTCTCCAGAAATTTTC 2073
DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY	2074	TCCTTTTGAAGTACTACTGATTTTACATTTGATAGGATGCTCTACAAGCTCATGATCTA 2133
DB	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY	2134	CAGCTCGAAAGAAATATCTCTACTGTCTGCTTTCATATATGTTGGTCTCTCAGGTGAGTTG 2193
DB	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660







Db 541 TyrValAenProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560  
QY 1894 ATCAGTCAGCAGCTGTGACTCTTTTATAAGTAAGTAGTAACACAGAGAAATCCACACTGT 1953  
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
QY 1954 GTGTCCTCTTTTAAAGCTATCAAGTCCTGAGAGATGACCCCACTGTCGAAACAAAGAAATTT 2013  
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
QY 2014 TGGGCCACCATTTTGGATTCCAGCAGGTCCTCTCTCGATATACCTCTCCGAAATTTTC 2073  
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620  
QY 2074 TCTTTTGAAGTACTACTGATTTACATTTGTATGGAGTCTCTCAAGCCCTCATGATCTA 2133  
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
QY 2134 CAGCCTCGAAAGAAATCTACTGTCTGTTCATATATGGTGGFCCTCAGGTGSCAGTTG 2193  
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660  
QY 2194 GTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTTGAATACCTAGCTCTCTAGGT 2253  
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGly 680  
QY 2254 TATGTGTTGTATGTATAGACACACAGGGATCCTGTCCAGGGGCTTAAATTTGAAGGC 2313  
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700  
QY 2314 GCCTTTAAATPATAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATAT 2373  
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720  
QY 2374 CTAGCTTCTCGATATGATTTCACTTAGATCGTGTGGCATCCACGGCTGCTCCTAT 2433  
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPsrTyr 740  
QY 2434 GGAGGATACCTCTCCCTGATGTCATTAATGACAGAGTTCAGATATCTTCAGGGTTCGTTT 2493  
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760  
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATCATACAGATACAGGACGGAAGTTATATG 2553  
Db 761 AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMet 780  
QY 2554 GGTCACTCGACAGAAAGACAGGGCTATTACTTAGGATCTGTGGCCATGCACAGAA 2613  
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800  
QY 2614 AGTTCCCTCTGACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTC 2673  
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820  
QY 2674 CATTTTGCATACCATATATTTACTGAGTTTCTGAGGCTGGAAAGCCATATGAT 2733  
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840  
QY 2734 TTACAGATCTATCTCCAGAGAGACACAGCATAAAGATTCCTGAATCCGGGAAACATTAT 2793  
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860  
QY 2794 GNACTGATCTTTGGCATCTTCAAGAAAACTTGGATCATCGTATTGCTGCTCTAAAA 2853  
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880  
QY 2854 GTGATA 2859  
Db 881 ValIle 882

RESULT 7  
ABU07720

ID XX ABU07720 standard; protein; 882 AA.  
AC ABU07720;  
XX  
DT 19-MAY-2003 (first entry)  
XX  
DE Human serine protease HIPHUM46.  
XX  
KW Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;  
KW serine protease activity modulation; dipeptidyl peptidase activity;  
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;  
KW Alzheimer's disease; paraspranuclear palsy; Huntington's disease;  
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;  
KW irritable bowel syndrome; type I diabetes; faecal incontinence;  
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;  
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;  
KW multiple sclerosis.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Region 259..260  
FT /note= "Paired glutamates of the beta propeller domain"  
FT Active-site 739  
FT /label= Catalytic\_serine\_residue  
FT Active-site 817  
FT /label= Catalytic\_aspartate\_residue  
FT Active-site 849  
FT /label= Catalytic\_histidine\_residue  
XX  
PN GB2374869-A.  
XX  
PD 30-OCT-2002.  
XX  
PR 22-JAN-2002; 2002GB-00001404.  
XX  
PR 23-JAN-2001; 2001GB-00001760.  
XX  
PA (GLAX ) GLAXO GROUP LTD.  
XX  
FI Edbrooke MR, Lewis AP;  
XX  
DR WFI; 2003-150703/15.  
DR N-PSDB; ABX12255.  
XX  
CC Identifying modulators of serine protease activity useful for treating  
CC musculoskeletal diseases, by contacting cell expressing a novel serine  
CC protease polypeptide with a compound and monitoring serine protease  
CC activity.  
PS Claim 10; Page 26-29; 38pp; English.  
XX  
CC The invention relates to a method of identifying a substance that  
CC modulates serine protease activity, comprising contacting a cell such as  
CC a neuronal cell, lung cell, intestinal cell or a cell infected with a  
CC virus, expressing a serine protease polypeptide (HIPHUM 46), or its  
CC variant having dipeptidyl peptidase activity, or a serine protease  
CC isolated from the cell with a test substance, and monitoring for serine  
CC protease activity. The method is useful for identifying a substance that  
CC modulates serine protease activity. A modulator of the serine protease is  
CC useful in the manufacture of a medicament for treatment or prophylaxis of  
CC a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus  
CC infection, Alzheimer's disease, paraspranuclear palsy, myotonic  
CC dystrophy, Huntington's disease, or amyotrophic lateral sclerosis.  
CC Additional disease that may be treated using modulators of the serine  
CC protease include malabsorption syndromes, irritable bowel syndrome, lung  
CC disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,  
CC rectal polyps, small bowel tumours, colorectal tumours, anaemia,  
CC dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple  
CC sclerosis. The present sequence represents the amino acid sequence of the  
CC human serine protease HIPHUM46  
XX  
SQ Sequence 882 AA;

Alignment Score:		0	Length: 882
Pred. No.:	882.00	882	Matches: 882
Score:	100.0%	0	Conservative: 0
Percent Similarity:	100.0%	0	Mismatches: 0
Best Local Similarity:	100.0%	0	Indels: 0
Query Match:	85.6%	0	Gaps: 0
DB:	6		
US-10-825-632-2 (1-3120) x ABU07720 (1-882)			
QY	214	ATGCGAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGAAACTGCCGACTGT	273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys	20
QY	274	GAGGAGAAATTTGAATCACAGGATCGGCCTAAATTGGAGCCCTTTTATGTGGCGGTAT	333
DB	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAAAGCTGTTCGCGATACAGAAATATCATGGCTCATGATG	393
DB	41	SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGACACATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATCA	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGATCTATTACCTTGCATGCTCGGTGAGACAGAGAAATACACTGTTTATTCT	513
DB	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAATTTCCAAAACATCAATAGCAGCAGCAGTCTTAATGCTCTCTTGTGAAAGCCCTTTTG	573
DB	101	GluLeuProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGGCAACACATGGACTATGGAATGTATTTCTCGAGAGAAGAACTATTAGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg	140
QY	634	GAAGAAGAACCATTTGAAACAGTCGGAAATTTCTTACGATATATCACAAGGAGTGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGAGGGGCCACAGGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATTTATCCCGCTGTATCCAGACTGGATGCTTTTATATACATAGCAACGATAT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATATCAATCATCGTAACACAGAGAAGAAGAGACTCCTATATGTGCACATGAGCTA	933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTTACCTTTGTTCTCCAGAAGAA	993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
QY	994	TTTGATAGATATTCTGCTATTGCTGTGTCCTCAAGAGCTGAAACACTCCAGTGGTGGT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGTGGAAATATTATCATGTT	1113
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAAACAGGAGGCGCAGATTCCTCGTTATCTCTAAACAGGTACA	1173

DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCTAAAGTCACTTTTAAAGATGTGAGAAATAATGATTTGATGCTGGAAGGAGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGCTCATAGATAAGGAACATAATCAACCTTTTGGAGATTTATTTGAAGGAGTTGAA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGCTGGAGTCTCTGAGGAAAAATATGCTTGTCTCCTACTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGAGCTCGCTACAGATAGTGTGATCTCACCTGATTAATTTATTTATTTATTTAT	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATGGAAAGGAGAGACTCATTGAGTCAGTCCCTGATTTCTGTGACGCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTetPleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAAGTCACGAAGGAAATTTAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAAATTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGGCTGCTCTCCAGTGATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATATACCACT	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GTCGAATGGGAAGTCTTTGGCCGCATGGATCTTAATATATCAAGTTTGATGAAGTCAGAAG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGACATCACCCTGACGTAGTCACT	1833
DB	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
QY	1834	TACGTAAATCTTGAGAGGTGACAAGCTGACCTGACCTGGCTGCTACTCACAATCTTGCTGC	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTGATCTTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGT	1953
DB	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGCTCCCTTTACAGCTATCAAGTCTGAGAGTACCCCAACTTGCAAAACAAAGGAATTT	2013
DB	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTCTGACTATATCTCTCCAGAAATTTTC	2073
DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
QY	2074	TCTTTTGAAGTACTACTGGATTTTACATTTATGGAATGCTCTTACAGGCTCATGATCTA	2133
DB	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCTCGAAAGAAATATCTTACTGCTGCTTCTATATATGCTGCTGCTGCTGCTGCTGCT	2193
DB	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu	660
QY	2194	GTGAATATCTGGTTAAAGGAGTCAGTATTTCCGCTTGAATACCTAGCCCTCTCTAGT	2253
DB	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680



QY	634	GAAGAAAAAGCATTGGACACAGTCGGAATTCCTTTCAGATTATCACCAAGAGAGTGA	693	Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	QY	1774	CTGTATATTTTGAAGGACCCAAAGACTCCCTTTTAGAGCATCACCTGTAGTACGTACGT	1833
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGATTTTACCGTAAAGATGAGGGCCACACAGGA	753	Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSer	540
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	QY	1834	TACGTAATCTCGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1893
QY	754	TTTACCAACAACCTTTAAGGCCCAATCTAGTGAATACTAGTTGTCCCAACATACGATG	813	Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	QY	1894	ATCAGTCACACTGTGACTTCTTTATAGTAAGTATAGTAACACAGAGAATCCACACTGT	1953
QY	814	GATCCAAAATATTGCCCCCTCATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	873	Db	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	QY	1954	GTCTCCCTTTACAAGCTATCAAGTCTCTGAAGTGAACCCAACTTGCACAAAACAAGGAATT	2013
QY	874	TGGATATCTAAACATCGTAACACGAGAGAAAGGAGACTCACTATGTGCACATGACTA	933	Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	221	TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu	240	QY	2014	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCCTCAGAGAAATTTTC	2073
QY	934	GCCAAATGGAAGAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAAGAGAA	993	Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260	QY	2074	TCCTTTGAAAGTACTACTGGATTACATTGTATGGATGCTCTACAAGCCTCATGATCTA	2133
QY	994	TTTGTAGATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCACAGTGGT	1053	Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	QY	2134	CAGCTGGAAAGAAATATCTACTGCTGTCTTATATATATATATATATATATATATATAT	2193
QY	1054	AAAATCTTGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAATTTATTCATGTT	1113	Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	QY	2194	GTGAATAATCGGTTTAAAGAGTCAAGTATTTCGCTTGAATACCTAGCCTCTCTAGGT	2253
QY	1114	ACATCCCTATGTGAAACAAGGAGGCGAGATTCATTCCTGTTTCTTAAACAAGGTACA	1173	Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	QY	2254	TATGTGTTGTAGTATAGACAAACAGGGGATCCTGTACCGAGGCTTAAATTTGAAGGC	2313
QY	1174	GCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATC	1233	Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	QY	2314	GCCTTTAAATATAAATG 2331	
QY	1234	ATGATCTCATAGTAAGGAACTAATTCAACCTTTTGAATTTTGAAGGAGTTGAA	1293	Db	701	AlaPheLysTyrLysMet 706	
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	RESULT 9			
QY	1294	TATATCCAGAGCTGGATGGAATCCTCGAGGGAATAATGCTTGCTCATCTACTAGAT	1353	ABU92026			
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	ID ABU92026	standard; protein; 831 AA.		
QY	1354	CGCTCCAGACTCGCTTACAGATAGTGTGATCTACCTGAAATTTATCCAGTAGAA	1413	XX AC	ABU92026;		
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	XX XX	15-JUL-2003 (first entry)		
QY	1414	GATGATTTATGGAAGGACAGACTCATTGAGTCAGTGCCTGATTCGTGACGCCACTA	1473	DT XX	Human protein modification and maintenance molecule-6 (PMM-6).		
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	DE XX	Human; protein modification and maintenance molecule; PMM; cancer;		
QY	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT	1533	XX KW	cell proliferation disorder; atherosclerosis; neurological disorder;		
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	XX KW	epilepsy; Huntington's disease; stroke; immune disorder; allergy;		
QY	1534	CCCCAAGTCAAGAGGAAATTTGAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTC	1593	XX KW	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;		
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	XX KW	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;		
QY	1594	CGTCATTTTATCAAAATTTACATCTATTTTTAAAGGAAGCAATATAACGATCCAGTGT	1653	XX KW	infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;		
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	XX KW	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary;		
QY	1654	GGGCTGCTGCTCCAGTGATTTCAAGTTCCTTCAAGAGGATAGCAATACCAAGT	1713	OS XX	antiinflammatory; thymimetic.		
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	OS OS	Homo sapiens.		
QY	1714	GGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATATCAAGTTGATGAAGTCAGAGG	1773	PN XX	WO2003031939-A2.		

```
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SX;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hatfield AJ, Chawla NK, Lehr-Wason BW, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang Y, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX WPI: 2003-430274/40.
DR N-PSDB; ACA92421.
XX
XX New human protein modification and maintenance molecules (PMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or
XX infections.
XX
XX Claim 1; Page 242-243; 31pp; English.
XX
XX The present invention relates to the isolation of human protein
XX modification and maintenance molecules (PMM), and the polynucleotide
XX sequences encoding them. A total of 40 PMM polypeptides (designated PMM
CC -1 to PMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders and infections. The
CC PMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
XX ABU92021-ABU92060 represent the human PMM polypeptides of the invention
XX
XX SQ Sequence 831 AA;
XX
XX Alignment Scores:
XX Pred. No.: 0 Length: 831
XX Score: 706.00 Matches: 706
XX Percent Similarity: 100.0% Conservative: 0
XX Best Local Similarity: 100.0% Mismatches: 0
XX Query Match: 68.5% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-10-825-632-2 (1-3120) x ABU92026 (1-831)
XX
XX 214 ATGGCAGCAGCAATCGAACAACAGACAGCTGGGTGTTGAGATATTTGAAACTCGGACTGT 273
XX 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
XX
XX 274 GAGGAGAATATTGAATACAGGATCGGCTAAATTTGGAGCTTTTATGTTGAGCGGTAT 333
XX 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
XX
XX 334 TCCTGGAGTCAGCTTAAAGAGCTCTTGGCGATACCAGAAAATATCATCGCTACATGATG 393
XX 41 SerTyrSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
XX
XX 394 GCTAAGGCACACATGATTTTCATGTTGTGAAGAGGAATGCCAGATGGACCTTCATCA 453
XX 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
XX
XX 454 GACGAATCTATTACCTTGGCTGTCGGTGAGACAGAAAATACACTGTTTATTCTTCT 513
XX 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrPheTyrSer 100
XX
XX 514 GAAATTCCTCCAAAACTATCAATAGACGACGAGCTTAAATGCTCTCTTGGAAAGCTCTTTTG 573
XX 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
XX
XX 574 GATCTTTTTCAGGCACACACTGGACTATGGAATGTATTTCTCGAAGAAGAAGAACTATTAGA 633
XX 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
XX 634 GAAAGAAAAAGCATTTGGAAACAGTCGGAATTTGCTTACGATTATCACCAGGAGGATGGA 693
XX 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
XX
XX 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCAGTAAAGAGTGGAGGCCACAGGA 753
XX 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
XX 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCTCCCAACATCAGGATG 813
XX 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
XX
XX 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
XX 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
XX 874 TGGATATCTAACATCTGTAACCCAGAGAAAGGAGACTCACTATGTGCAATGACCTA 933
XX 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
XX 934 GCCAATCATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAA 993
XX 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
XX 994 TTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCACAGTGGTGT 1053
XX 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
XX 1054 AAAATTTCTTAGAATCTTATATGAAGAAATCATCAATCTCAGGTGGAAATATTTCATGTT 1113
XX 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
XX 1114 ACATCCCTTATCTTGGAAACCAAGGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTACA 1173
XX 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
XX 1174 GCAAAATCTCTAAAGTCACCTTTTAAAGATGTGAGAAATATGATGCTCGAAGGAAGATC 1233
XX 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
XX 1234 ATAGATGTGCATAGATAAGAACTAATTCACCTTTTGAGATTCCTATTTGAGGAGCTTGA 1293
XX 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
XX
XX 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTTGGTCCATCTCTACTAGAT 1353
XX 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
XX 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCACTGAATATTATTTCCAGTAGAA 1413
XX 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
XX 1414 GATGATGTTTGAAGAGGAGAGACTCATGAGTCAGTCGCTGATTCGTGTCAGCCACTA 1473
XX 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
XX 1474 ATTATCTATGAAGAAAACAAAGACATCTGGATAAATATCATGACATCTTTTCATGTTTT 1533
XX 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
XX 1534 CCCCAGATCTCAGAGAGGAAATTTGAGTTTATTTTTCGCTCTGGAATGCAAAAACAGGTTTC 1593
XX 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
XX 1594 CGTCATTTATACAAATTTACATCTCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGT 1653
```



Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
QY 1654 GGGTGGCTGCTCAAGTATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATACCAGT 1713  
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleGalaIleThrSer 500  
QY 1714 GGTGAATGGGAAGTTCTTGGCCGCATGATCTTAATATCAAGTTGATGAAGTCAAGG 1773  
Db 501 GlyGluTrpGluValLeuGlyA-GHisGlySerAsnIleGlnValAspGluValArgArg 520  
QY 1774 CTGGTATATTTTGAAGGACAAAGACTCCCTTTAGACATCACTGTACTGACTGCT 1833  
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540  
QY 1834 TACGTAATCTCGAGAGGTGACAGGCTGACGCTGACGCTGCTTCACTTCTGCTGTC 1893  
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560  
QY 1894 ATCAGTACGACTGTGACTCTCTTTATAGTAAGTATAGTAACAGAGAATCCACACTGT 1953  
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
QY 1954 GTGTCCCTTTTACAAGCTATCAAGTCTTGAAGATGACCCACTGTGAAACAAAGCAATT 2013  
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
QY 2014 TGGGCCACATTTGGATGACGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073  
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspPyrThrProGluIlePhe 620  
QY 2074 TCTTTTGAAGTACTTACTGGATTACATTGTATGGATGCTCTACAGGCTCATGATCTA 2133  
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
QY 2134 CAGCTTGGAAAGAAATATCTACTGTGCTGTTCTATATATGTTGCTCTGAGTGCAGTTG 2193  
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660  
QY 2194 GTGAATACGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTTAGT 2253  
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680  
QY 2254 TATGTGGTGTAGTGATGACACACAGGAGTCTCTGTACCCGAGGCTTAAATTTGAAGC 2313  
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700  
QY 2314 GCCTTTAAATATAAATG 2331  
Db 701 AlaPheLysTyrLysMet 706  
RESULT 10  
ID ABU92030 standard; protein; 824 AA.  
XX AC ABU92030;  
XX DT 15-JUL-2003 (first entry)  
XX DE Human protein modification and maintenance molecule-10 (PMM-10).  
XX KW Human; protein modification and maintenance molecule; PMM; cancer;  
KW cell proliferation disorder; atherosclerosis; neurological disorder;  
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;  
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;  
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;  
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary;  
KW antiinflammatory; thyromimetic.  
XX OS Homo sapiens.  
XX PN WO2003031939-A2.

XX PD 17-APR-2003.  
XX PF 11-OCT-2002; 2002WO-US032850.  
XX PR 12-OCT-2001; 2001US-0329689P.  
XX PR 25-OCT-2001; 2001US-0335703P.  
XX PR 09-NOV-2001; 2001US-0348887P.  
XX PR 28-NOV-2001; 2001US-0334145P.  
XX PR 06-DEC-2001; 2001US-0337451P.  
XX PR 14-DEC-2001; 2001US-0340584P.  
XX PA (INCYTE GENOMICS INC.  
XX PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
PI Tran UK, Becha SD, Duggan BW, Lee EA, Griffin JA, Li JX;  
PI Sprague WP, Hatalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;  
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;  
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
XX MPI; 2003-430274/40.  
DR N-PSDB; ACA92425.  
XX New human protein modification and maintenance molecules (PMM), useful  
PT for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or  
PT infections.  
XX Claim 1; Page 249-251; 31pp; English.  
XX The present invention relates to the isolation of human protein  
XX modification and maintenance molecules (PMM), and the polynucleotide  
XX sequences encoding them. A total of 40 PMM polypeptides (designated PMM  
XX -1 to PMM-40) are disclosed. The sequences of the invention are useful  
XX for diagnosing a condition or disease associated with the expression of  
XX PMM in a subject, preparing a polyclonal or monoclonal antibody, and  
XX generating an expression profile of a sample containing the  
XX polynucleotides. The diseases or conditions associated with decreased  
XX expression or overexpression of PMM are cell proliferation disorders  
XX (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
XX Huntington's disease, stroke), immune/inflammatory disorders (e.g. AIDS,  
XX allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
XX syndrome), gastrointestinal or epithelial disorders, and infections. The  
XX PMM polypeptides or their fragments are useful in screening compounds  
XX for effectiveness as agonists or antagonists of the polypeptides, or in  
XX altering the expression of the target polynucleotide and compounds that  
XX specifically bind to, or modulate the activity of the polypeptide.  
XX ABU92021-ABU92060 represent the human PMM polypeptides of the invention  
SQ Sequence 824 AA;  
Alignment Scores:  
Pred. No.: 0 Length: 824  
Score: 700.00 Matches: 700  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 68.0% Indels: 0  
DB: Gaps: 0  
US-10-825-632-2 (1-3120) x ABU92030 (1-824)  
QY 760 CAACACCTTTAAGGCCCAATCTAGTGGAACTAGTGTGTCACATACCGATGATCCA 819  
Db 125 GlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAspPro 144  
QY 820 AAATATGCCCCGCTGATCCAGACTCGATGCTTTTATATACATAGCAACGATTTGGATA 879  
Db 145 LysLeuCysProAlaAspProAspTPIleAlaPheIleHisSerAsnAspIleTrpIle 164  
QY 880 TCTAATCTGTAAACAG 939  
Db 165 SerAsnIleValThrArgGluArgGluArgGluArgGluArgGluArgGluArgGlu 184

QY 940 ATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCTCAAGAGAATTTGAT 999  
Db 185 MetGluGluAaspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAasp 204  
QY 1000 AGATATCTTGCTTATGCTGGTGTCCAAAAGCTGAAACAACCTCCAGAGTGGTGAATAAT 1059  
Db 205 ArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGlyVal 224  
QY 1060 CTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTTCATGTTACATCC 1119  
Db 225 LeuArgIleLeuTyrGluGluAenAaspGluSerGluValGluIleIleHisValThrSer 244  
QY 1120 CCTATGTGGAACACAGGAGGCGAGATTCATTCCGCTTATCTCTAAACNGGTACAGAAAT 1179  
Db 245 ProMetLeuGluThrArgAlaAaspSerPheArgTyrProLysThrGlyThrAlaAen 264  
QY 1180 CCTAAAGTCACTTTTAAGATGTCAGAAATATGATGTCGTGAAGGAGATCATAGAT 1239  
Db 265 ProLysValThrPheLysMetSerGluIleMetIleAepAlaGluGlyArgIleAep 284  
QY 1240 GTCATAGATGAAGAACTAATCAACCTTTTGAGATCTATTGGAAGGAGTTGAATATTT 1299  
Db 285 ValIleAaspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIle 304  
QY 1300 GCCAGAGCTGGATGCTCTCAGGAGAAATATATGCTGGTCCATCTCTACTAGATCGCTCC 1359  
Db 305 AlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAaspArgSer 324  
QY 1360 CAGACTCGCTACAGATAGTGTGATCTCACTCAATTTATTTGAGAGGAGTTGAATATTT 1419  
Db 325 GlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAaspAep 344  
QY 1420 GTTATGGAAGCAGAGACTCATTTAGTCAGTGCCTGATTTCTGTGACGCCACTAATTATC 1479  
Db 345 ValMetGluArgGlnArgLeuIleGluSerValProAepSerValThrProLeuIleIle 364  
QY 1480 TATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTCCCAA 1539  
Db 365 TyrGluGluThrThrAspIleTrpIleAsnIleHisAaspIlePheHisValPheProGln 384  
QY 1540 AGTCACGAGAGAAATGAGTTTATTTTGGCTCTGMAATGCAACACAGTTCCGTCAT 1599  
Db 385 SerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHis 404  
QY 1600 TTATCAAAATTTACATCTTATTTAAAGGAAGCAATATATAACGATCCAGTGGTGGCTG 1659  
Db 405 LeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLysLeu 424  
QY 1660 CCTGCTCCAAGTATTTCAAGTGTCTATCAAGAGGAGATAGCAATTCACGTCGTGAA 1719  
Db 425 ProAlaProSerAaspPheLysCysProIleLysGluIleAlaIleThrSerGlyGlu 444  
QY 1720 TGGGAAGTTCTTGGCCGCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGCTGTA 1779  
Db 445 TrpGluValLeuGlyArgHisGlySerAsnIleGlnValAaspGluValArgArgLeuVal 464  
QY 1780 TATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAGTTACGTA 1839  
Db 465 TyrPheGluGlyThrLysAaspSerProLeuGluHisLeuTyrValValSerTyrVal 484  
QY 1840 AATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGT 1899  
Db 485 AsnProGluValThrArgLeuThrAepArgGlyTyrSerHisSerCysCysIleSer 504  
QY 1900 CAGCACTGCTACTCTTTATAGTATAGTATACCAAGAGAATCCACACTGTGTGCTCC 1959  
Db 505 GlnHisCysAaspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSer 524  
QY 1960 CTTTACAGCTATCAAGTCTCAAGATGACCCAACTTCGCAAAACAAAGGAATTTGGGCC 2019  
Db 525 LeuTyrLysLeuSerSerProGluAaspAepProThrCysLysThrLysGluPheThrPala 544  
QY 2020 ACCATTTTGGATTCAGCAGGTCTCTCTCTGACTATACTCTCCAGAAATTTTCTCTTTT 2079

Db 545 ThrIleLeuAaspSerAlaGlyProLeuProAaspTyrThrProProGluIlePheSerPhe 564  
QY 2080 GAAAGTACTACTTGATTTACATTTGATGCGATGCTCTACAAGCCCTCATGATCTACAGCT 2139  
Db 565 GluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAepLeuGlnPro 584  
QY 2140 GGAAAGAAATATCTACTGTCTGCTTATATATGATGCTGCTCTCAGGTGCAGTTCGTGAAT 2199  
Db 585 GlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsn 604  
QY 2200 AATCGGTTTAAAGAGCTCAAGTATTTCCGCTTGATACCTAGCTCTCTAGGTATGTG 2259  
Db 605 AsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal 624  
QY 2260 GTTGTAGTATGACACAACAGGGATCTGTCACCGAGGCTTAAATTTGAAGCGCCTTT 2319  
Db 625 ValValValIleAaspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPhe 644  
QY 2320 AAATATAAATGGGTCAATAGAAATTCAGCATCAGGTGAGGACTCCCAATATCTAGCT 2379  
Db 645 LysTyrLysMetGlyGlnIleGluIleAaspAepGlnValGluGlyLeuGlnTyrLeuAla 664  
QY 2380 TCTCGATATGATTTCAATTCAGCTTAGATCTGTGGGCATCCACGGCTGGTCTTATGGAGGA 2439  
Db 665 SerArgTyrAspPheIleAaspLeuAaspArgValGlyIleHisGlyTrpSerTyrGlyGly 684  
QY 2440 TACCTCTCCCTGATCGCATTAATGCAGAGTCAGATATCTCAGGGTTGCTATTGCTGGG 2499  
Db 685 TyrLeuSerLeuMetAlaLeuMetGlnArgSerAaspIlePheArgValAlaIleAlaGly 704  
QY 2500 GCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATATGGTGCAC 2559  
Db 705 AlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHis 724  
QY 2560 CTGTACCAAGAAATCAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTC 2619  
Db 725 ProAepGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhe 744  
QY 2620 CCTCTGAACAAATCGTTTACTGCTCTTACATGGTTTCTGATGAGAAATGTCCATTTT 2679  
Db 745 ProSerGluProAenArgLeuLeuLeuHisGlyPheLeuAaspGluAenValHisPhe 764  
QY 2680 GCACATACCACTATATTACTGAGTTTATAGTCAGGGCTGGAAAGCCATATGATTTACAG 2739  
Db 765 AlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAaspLeuGln 784  
QY 2740 ATCTATCTCAGGAGACACACAGATAAGAGTTCTCTGAATCGGAGAACATTTACACTG 2799  
Db 785 IleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGlnLeu 804  
QY 2800 CATCTTTGCACTACCTTCAAGAAACCTTGGATCACGTTATGCTGCTCTAAAGTGTATA 2859  
Db 805 HisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaLeuLysValIle 824  
RESULT 11  
ABB97361  
ID ABB97361 standard; protein; 782 AA.  
XX ABB97361;  
AC ABB97361;  
DT 27-JUN-2002 (first entry)  
XX  
DE Novel human protein SEQ ID NO: 629.  
XX  
KW Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;  
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
XX expressed sequence tag.  
XX Homo sapiens.  
XX  
XX WO200222660-A2.  
FN

XX 21-MAR-2003.  
PD  
XX  
PF 10-SEP-2001; 2001MO-US026015.  
XX  
PR 11-SEP-2000; 2000US-00659671.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Zhou P, Auendi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX  
XX WPI: 2002-292408/33.  
DR N-PSDB; ABN32547.  
XX  
PT An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis.  
XX  
PS Example 2; SEQ ID NO 629; 509pp; English.  
XX  
CC The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke  
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.  
CC rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a protein of the invention  
XX  
SQ Sequence 782 AA;  
  
Alignment Scores:  
Pred. No.: 0 Length: 782  
Score: 658.00 Matches: 658  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 63.9% Indels: 0  
DB: 5 Gaps: 0  
  
US-10-825-632-2 (1-3120) x ABB97361 (1-782)  
  
QY 214 ATGCAGCAGCAATGGAACAGAACAGCTGGGTGGATGATTTGAAATCGCGACTGT 273  
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
  
QY 274 GAGGAGATATTGAATCACAGGATCGCCTAAATTGGAGCCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
  
QY 334 TCCTGGAGTCACTTAAAGCTTGTCCGATACCAGAAATATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
  
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCAITCA 453  
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
  
QY 454 GACAGAATCTATTACCTTGGCATGTCCTGGTGAACACAGAGAAATATACACTGTTTATTC 513  
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
  
QY 514 GAAATTCCTCAAAATATCAATAGACGACGATCTTAATGCTCTCTTGAAGCCTCTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
  
QY 574 GATCTTTTCAGGCAACACTGGACTATGGATGATATCTTCGAGAGAGAAGAACTATTAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
  
QY 634 GAAAGAAACCATTTGAAACAGTCGGAATTCCTTACGATATATCACAAGGAAGTGA 693  
Db  
  
141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY  
694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCATCGTAAAGATGGAGGCCCAAGGA 753  
Db  
161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY  
754 TTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTGTCCTCCAAACATACGATG 813  
Db  
181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY  
814 GATCCAAAATTTATGCCCCGCTGATCCAGACTGGATGCTTTTATATACATAGCAACGATATT 873  
Db  
201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY  
874 TGGATATCTAACAATCGTAACACAGAGAAGAAAGAGACTCATTATGTGTCACATAGAGCTA 933  
Db  
221 TrpIleSerAsnIleValThrArgGluGluArgGLeuThrTyrValHisAsnGluLeu 240  
QY  
934 GCCAACATGGAAGAGATGCCAGATCAGCTGAGTGGCTGCTTACCTTTGTTCTCCAGAAGAA 993  
Db  
241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260  
QY  
994 TTTGATAGATATCTGCTGCTATTGCTGCTGCTCAAAAGCTGAAACAACTCCCAAGTGGTGGT 1053  
Db  
261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY  
1054 AAAATTTCTTGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAATATTATTTCATGTT 1113  
Db  
281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300  
QY  
1114 ACATCCCTCATGTTGGAACAACAGGAGGCGAGATTTCATTCGCTTATCTCTAAACAGGTACA 1173  
Db  
301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY  
1174 GCAAAATCCTAAAGTCACTTTTAAGATGTGAGAAATAATGATTGATGCTGGAAGGAAGATC 1233  
Db  
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY  
1234 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGGAGATTTCTATTGAGGAGTTCAA 1293  
Db  
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
QY  
1294 TATATTCGACAGCTGGATGGACTCCTGAGGGAATAATATGCTTGTGCTCATCTACTAGAT 1353  
Db  
361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380  
QY  
1354 CGCTCCACAGACTCGCTTACAGATAGTGTGATCTCACCTGGAATATTATTATCCAGTAGAA 1413  
Db  
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY  
1414 GATGATGTTATGGAAGGACGACAGACTCATGAGTCAGTCCCTGATCTTGTGACGCCACTA 1473  
Db  
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY  
1474 ATTATCTATGAAGAAACAACAGACATCTGGAATAAATATCCATGACATCTTTCATGTTTTT 1533  
Db  
421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
QY  
1534 CCCCAAAGTCCAGNAGGAAATGATTTATTTTGGCTCTGATGATCAAAACAGAGTTTC 1593  
Db  
441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
QY  
1594 CGTCATTTATACAAAATATCATCTATTTTAAAGGAAGCAATATAAACGATCCAGTGGT 1653  
Db  
461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
QY  
1654 GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTTACCACT 1713  
Db  
481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500  
QY  
1714 GGTGAATGGGAAGTCTTGTGGCCGCATGATCTTAATATCAAGTTGTGTAAGTCAGAGG 1773  
Db  
501 GlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGluValArgArg 520



Db	601	TtpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
Qy	2074	TCCTTTTGAAGTACTACTGGATTTACATGTTATGGGATGCTCTACAAGCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	2134	CAGCTCGGAAGAAATATCCCTACTGCTGCTTTCATATATGTTGGT	2178
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly	655
RESULT 13			
ABG61596			
ID	ABG61596 standard; protein; 661 AA.		
XX	ABG61596;		
DT	12-AUG-2002 (first entry)		
DE	Human DPPP-1 splice variant #3.		
XX	Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;		
KW	DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;		
KW	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;		
KW	heart failure; hypertension; urinary retention; osteoporosis; cancer;		
KW	ulcer; allergy; cancer; psychotic disorder; neurological disorder;		
KW	dyskinesia; reproductive disorder; inflammatory disorder;		
KW	metabolic disorder.		
OS	Homo sapiens.		
XX	WO200231134-A2.		
XX	18-APR-2002.		
PF	12-OCT-2001; 2001WO-US031874.		
XX	12-OCT-2000; 2000US-0240117P.		
PR	(FERR ) FERRING BV.		
PA	Qi S, Akinsanya KO, Riviere PJ, Junien J;		
PI	WPI; 2002-444178/47.		
XX	N-PSDB; ABK83327.		
DR	New dipeptidyl peptidase IV-related proteins and nucleic acids encoding		
PT	the proteins, useful for treating e.g. fungal, bacterial, protozoan and		
PT	viral infections, cancers, allergies, neurological disorders, or pain.		
XX	Disclosure; Page 63-65; 113pp; English.		
XX	The present invention relates to the isolation of novel human serine		
CC	proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins		
CC	(DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic		
CC	acids encoding them are useful for treating infections such as fungal,		
CC	bacterial, protozoan and viral infections, particularly infections caused		
CC	by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,		
CC	precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's		
CC	disease, acute heart failure, hypotension, hypertension, urinary		
CC	retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,		
CC	allergies, cancers, migraine, vomiting, psychotic and neurological		
CC	disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.		
CC	These may also be used in discovering therapeutic agents for the		
CC	treatment of reproductive, inflammatory and metabolic disorders. ABG61591		
CC	-ABG61612 represent human DPPP proteins		
XX	Sequence 661"AA;		
SQ	Alignment Scores:		
Pred. No.:		0	Length: 661
Score:		655.00	Matches: 655
Percent Similarity:		100.0%	Conservative: 0
Best Local Similarity:		100.0%	Mismatches: 0

Query Match:	63.6%	Indels:	0
DB:	5	Gaps:	0
US-10-825-632-2 (1-3120) x ABG61596 (1-661)			
QY 214	ATGCGAGCAGCAATATGGAACACAGACCTGGGTGGTGGATATTTGAAACGCGACTGT	273	
DB	1 MetAlaAlaMetGluThrGluGlnLeuGluValGluLeuPheGluThrAlaAspCys	20	
QY 274	GAGAGAAATATTGAATCACAGGATCGGCTAAATTTGAGCCCTTTTATGCTGACGGTAT	333	
DB	21 GluGluAsnLeuGluSerGlnAspArgProLysLeuGluProPheArgValGluArgTyr	40	
QY 334	TCCTGGAGTCAGCTTAAAGCTCCTGCGATACAGAAATATCATGGCTCATCATG	393	
DB	41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60	
QY 394	GCTAAGCACCACATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTC	453	
DB	61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80	
QY 454	GACGAATCTATTACCTTGCATGCTGTGTGAGAACAGAGAAATATACACTGTTTATTC	513	
DB	81 AspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100	
QY 514	GAATTTCCCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTG	573	
DB	101 GluLeuProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu	120	
QY 574	GAFTCTTTTTCAGCAACACTGGACTATGGAATGATTTCTCGAGAGAGAACTATTAAAG	633	
DB	121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuLeuArg	140	
QY 634	GAAGAAGACGCAATGGAACTCGGAATTTGCTTACGATTTATCAACAAAGAACTGGA	693	
DB	141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	
QY 694	ACATTTCTGTTTCAAGCCCGTAGTGGAAATTTATCACGTAAAGATGAGGGCCACAGCA	753	
DB	161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	
QY 754	TTTACCCCAACACCTTTAAGGCCCAATCTAGTGAACACTAGTGTCCCAACATACGGATG	813	
DB	181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	
QY 814	GATCCAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATATACATAGCAACGATAT	873	
DB	201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	
QY 874	TGGATATCTAACATCGTAACAGAGAGAAAGGAGACTCACTATGTGCACAAATGAGCTA	933	
DB	221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	
QY 934	GCCACATGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTCTCCAAGAGAA	993	
DB	241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260	
QY 994	TTTGATAGATTCGTGCTATTGTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGT	1053	
DB	261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly	280	
QY 1054	AAATTTCTAGAATTTATATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTT	1113	
DB	281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	
QY 1114	ACATCCCTATGTTGGAAACAGGAGGCGAGATTCATTCCGTTTATCTTAAACAGGTACA	1173	
DB	301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	
QY 1174	GCAATCCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGAGGAGGATC	1233	
DB	321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	

QY 1234	ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGGATGTTCTATTTGAAGAGTTGAA	1293	
DB	341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	
QY 1294	TATATTGCCAGAGCTGGATGAGTCTCTGAGGGAATAATATGCTTGGTCCATCTCTAGAT	1353	
DB	361 TyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	
QY 1354	CGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATATTTATCCCAAGTAGAA	1413	
DB	381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	
QY 1414	GATCATGTTATGAAAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTCACCCACTA	1473	
DB	401 AspAspValMetGluArgGlnLeuIleGluSerValProAspSerValThrProLeu	420	
QY 1474	ATTATCTATCAAGAAACACACAGACTCTGGATAAATATCCATGACATCTTCATGTTT	1533	
DB	421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	
QY 1534	CCCCAAGTCACGAAGAGAAATTCAGTTTATTTTGCCTCTGCTGAAATCGAAAACAGTTTC	1593	
DB	441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	
QY 1594	CGTCATTTATACAAAATTCATCTATTTTAAAGGAAAGCAAAATATAAACCATCCAGTGT	1653	
DB	461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	
QY 1654	GGCTCCCTGCTCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTTACAGT	1713	
DB	481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500	
QY 1714	GGTGAATGGGAATGTTCTTGGCGGATCGATCTTAATATCCAAAGTTGATGAAGTCAGAAG	1773	
DB	501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	520	
QY 1774	CTGGTATATTTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCA	1833	
DB	521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540	
QY 1834	TACGTAATTCCTGGAGAGTGCACCAAGGCTGACTCACCGTGGCTACTCACATTTCTGCTGC	1893	
DB	541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	
QY 1894	ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAGAAATCCACACTCT	1953	
DB	561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	
QY 1954	GTGTCCCTTTTACAAGCTATCAAGTCCCTGAAGATGACCCCACTTGCACCAAGGAATTT	2013	
DB	581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	
QY 2014	TGGGCCACATTTTGGATTACAGAGGCTCTTCTTCCCTGACTATCTCTCCAGAAATTTTC	2073	
DB	601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	
QY 2074	TCCTTTGAAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAGGCTCATGATCTA	2133	
DB	621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	
QY 2134	CAGCTCGAAGAAATATCTACTGCTGCTTTCATATATGTTGGT 2178		
DB	641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655		

## RESULT 14

ABG61594  
ID ABG61594 standard; protein; 690 AA.  
XX  
XX ABG61594;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE Human DPRP-1 splice variant #1.



XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;  
XX DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;  
KW Diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;  
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;  
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;  
KW dyskinesia; reproductive disorder; inflammatory disorder;  
KW metabolic disorder.  
XX Homo sapiens.  
XX W0200231134-A2.  
XX 18-APR-2002.  
XX 12-OCT-2001; 2001WO-US0311874.  
XX 12-OCT-2000; 2000US-0240117P.  
XX (FERR ) FERRING BV.  
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;  
XX WPI; 2002-444178/47.  
XX N-PSDB; ABK83325.  
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding  
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and  
PT viral infections, cancers, allergies, neurological disorders, or pain.  
XX Disclosure; Page 59-61; 113pp; English.  
XX The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins  
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic  
CC acids encoding them are useful for treating infections such as fungal,  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591  
CC -ABG61612 represent human DPPP proteins  
XX Sequence 690 AA;  
SQ  
Alignment Scores:  
Pred. No.: 0 Length: 690  
Score: 655.00 Matches: 655  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 63.6% Indels: 0  
DB: 5 Gaps: 0  
US-10-825-632-2 (1-3120) x ABG61594 (1-690)  
QY 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGGTGGATATTTGAACTGCGGACTGT 273  
Db 1 MetAlaAlaMetGluThrGluGluLeuGluValGluLeuPheGluThrAlaAspCys 20  
QY 274 GNGGGAATATTGAATCACAGGATCGGCCTAAATTCGAGCCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAenIIeGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGTTGCCGATACAGAAAATATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 GCTAAGGCACCATGATTTTCATGTTGTGAAGGAATGATCCAGATGGACCTCATTC 453

Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 GACAGATCTATTACTCTGCCATCTCTGTCGAGAACAGAGAAATACACTGTTTATTCT 513  
Db 81 AspArgileTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
QY 514 GAAATTTCCAAACTATCAATAGACGACAGCTTAAATGCTCTCTTGGAAAGCCTCTTTG 573  
Db 101 GluileProLysThrileAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTCAGGCAACACTGGACTATGCAATGATTTCTCGAGAAGAAGAACTATTAA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
QY 634 GAAAGAAAACGCAATTGGAAACAGCTCGGAATTCCTTTACGATTATATCCAGGAAGTGA 693  
Db 141 GluArgLysArgileGlyThrValGlyLlealaserTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyileTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTGTCTCCCAACATACGGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnileArgMet 200  
QY 814 GATCCAAATATGCCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873  
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTAAACATCGTAAACCAGAGAAGAAAGGAGACTCACTATGTGCAATGAGCTA 933  
Db 221 TrpIleSerAsnileValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240  
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACTCTTGTCTCCAGAGAGAA 993  
Db 241 AlaAsnMetGluGluAspAlaahSerAlaGlyValAlaIleThrPheValLeuGlnGlu 260  
QY 994 TTTGATAGATATTCTCGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCACAGTGGT 1053  
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATTTCTAGAATTTCTATAGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113  
Db 281 LysileLeuArgileLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300  
QY 1114 ACATCCCTATGTTGAAACACAGAGGCGAGATTCACTCCGTTCCTTAAACACAGGTACA 1173  
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAATCTCAAGCTCACTTTTAAGATGTCAGAAATAATGATGATGCTGAAGGAAGATC 1233  
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGAGATTCTATTGAAGGAGTTGAA 1293  
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluileLeuPheGluGlyValGlu 360  
QY 1294 TATATTGCCAGAGCTGAGTGGACTCTCTGAGGGAATAATGCTGTGCTCCTACTAGAT 1353  
Db 361 TyrileAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerileLeuLeuAsp 380  
QY 1354 CGCTCCAGACTCGCTTACAGATAGTGTGATCTACCTGAAATTTATTTATCCAGTAGAA 1413  
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATGATGTTATGGAAGACGAGACTCATGTGAGTCAGTCGCTGATTTCTGTGACGCCACTA 1473  
Db 401 AspAspValMetGluArgGlnArgLeuileGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATGAAGAAACAAACAGACATCTGATTAATATCCATGACATCTTTTCATGTTT 1533  
Db 421 ileileTyrGluGluThrThrAspIleTrpIleAsnileHisAspIlePheHisValPhe 440



161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
754	TTTACGCCAACAACTTTAAGGCCCACTAGTAGGAACACTAGTTGTCGCCAACATACGGATG	813
181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
814	GATCCAAATATTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATACGACACGATATT	873
201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
874	TGCATATCTTAACATCGCTAAACCAGAGAAAGAGAGACTCATTATGTGTCACAAATGAGCTA	933
221	TrpIleSerAsnIleValThrArgGluLysArgLeuThrTyrValHisAsnGluLeu	240
934	GCCAACATCGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAA	993
241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
994	TTTGATAGATATTCTGGCTATTGTGTGTGTCCAAAGCTGAAACAACTCCAGTGGTGGT	1053
261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
1054	AAATTCCTTGAATTCCTATATGAAGAAATGATGAATCTGAGTGCAGAAATTTATTCATGTT	1113
281	LysIleLeuArgIleLeuTyrGluLysAsnAspGluSerGluValGluIleHisVal	300
1114	ACATCCCCTATGTTGGAACAAGAGGAGCGAGATTCAATCCGTTATCTCTAAACACAGTGACA	1173
301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
1174	GCAAATCCTTAAGTCACCTTTTAAGATGTCAGAAATTAATGATGCTCGAAGGAAGGATC	1233
321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
1234	ATAGATGTCATAGATAGAAGAACTAATTCAACCTTTTGAGATTCATTATTTGAAGAGATTGAA	1293
341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
1294	TATATTGCCAGAGCTGGATGACTCCTGAGGGAAATATGCTTGCTCCATCCTACTAGAT	1353
361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATTTATTTATCCCAAGTAGAA	1413
381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
1414	GATGATGTTATGGAAGGCAGAGACTCATTTAGCTCAGTCGCTGATTTCTGTGACGCCACTA	1473
401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
1474	ATTATCTATGAAGAAACAACAGACATCTGATTAATATCCATGACATCTTTCATGTTTTT	1533
421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
1534	CCCCAAGTCACGAAGAGAAATTTAGTTTATTTTTGCTCTGAATGCAAAACAGGTTTC	1593
441	ProGlnSerHisGlnGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
1594	CGTCATTTATCAAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1653
461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
1654	GGGCTCCCTGCTCCAAGTGATTTCAGTGTGCTTATCAAAGAGAGATAGCAATTTACCAGT	1713
481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
1714	GGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAAATATCCAAGTTGATGAAGTCAGAAGG	1773
501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
1774	CTGGTATATTTTGAAGGCCCAAAAGACTCCCTCTTTAGCGATCACCTGTACGTAGTCAGT	1833

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 03:26:34 ; Search time 20.8 Seconds  
(without alignments)  
2886.503 Million cell updates/sec

Title: US-10-825-632-2  
Perfect score: 1030  
Sequence: 1 aggtgtaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 60.0 , Fgapext 7.0  
Delop 60.0 , Delext 60.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 565918

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-O=/abss/ABSSWEB.spool/US10825632/runat\_01052006\_110237\_4459/app\_query.fasta\_1  
-DB=PIR -QFMT=fastan -SUFFIX=oligo.n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p  
-USER=US10825632 @CGN 1.1\_63 @runat\_01052006\_110237\_4459 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database :

PIR\_90:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	9	0.9	249	2	T09139	26S proteasome alp
C 2	9	0.9	297	2	H84168	molybdenum cofacto
C 3	9	0.9	667	2	S63587	gene pacC protein
C 4	9	0.9	721	2	F70649	probable Acyl-CoA
C 5	9	0.9	2073	1	EWASBE	bimE protein - Eme
C 6	9	0.9	2194	1	GNVYE7	genome polyprotein
C 7	8	0.8	65	2	AH2228	hypothetical prote
C 8	8	0.8	85	1	FDFL4W	antifreeze protein
C 9	8	0.8	91	2	A22592	antifreeze protein
C 10	8	0.8	97	2	S02376	antifreeze protein
C 11	8	0.8	113	2	B75632	hypothetical prote
C 12	8	0.8	123	2	F64083	ribosomal protein
C 13	8	0.8	131	1	SEPG	secretin precursor
C 14	8	0.8	141	2	F38164	hypothetical prote

15	8	0.8	145	2	AE2962	conserved hypothet
16	8	0.8	150	2	S09872	hypothetical prote
17	8	0.8	156	2	S77503	hypothetical prote
18	8	0.8	166	2	T49489	related to H+-tran
19	8	0.8	180	2	A98321	hypothetical 15.0K
20	8	0.8	183	2	A72657	VPS29-like phospho
21	8	0.8	200	2	GB4822	probable LIM-domai
22	8	0.8	203	2	JC6113	melanocyte-specifi
23	8	0.8	205	2	S70618	transcription regu
C 24	8	0.8	216	2	B26423	serine proteinase
C 25	8	0.8	216	2	A61259	glycoprotein S - p
C 26	8	0.8	225	1	MMIH68	E1 membrane glycop
C 27	8	0.8	244	2	T00449	hypothetical prote
C 28	8	0.8	247	2	AG3513	channel protein vi
C 29	8	0.8	250	2	C87443	conserved hypothet
C 30	8	0.8	264	2	E95410	probable ABC trans
C 31	8	0.8	275	2	S21348	probable pol polyp
C 32	8	0.8	301	2	H95308	probable ABC trans
C 33	8	0.8	303	2	D84348	hypothetical prote
C 34	8	0.8	308	2	C83586	probable transcrip
C 35	8	0.8	310	2	A60967	insulin-like growt
C 36	8	0.8	317	2	I46916	insulin-like growt
C 37	8	0.8	321	2	T12497	hypothetical prote
C 38	8	0.8	330	1	H69798	conserved hypothet
C 39	8	0.8	342	2	H70653	hypothetical prote
C 40	8	0.8	346	2	T13973	NADH2 dehydrogenas
C 41	8	0.8	346	2	T14033	NADH2 dehydrogenas
C 42	8	0.8	346	2	T14049	NADH2 dehydrogenas
C 43	8	0.8	346	2	T13835	NADH2 dehydrogenas
C 44	8	0.8	346	2	T17105	NADH2 dehydrogenas
C 45	8	0.8	346	2	T13834	NADH2 dehydrogenas

ALIGNMENTS

RESULT 1

T09139  
26S proteasome alpha chain - spinach  
C:Species: Spinacia oleracea (spinach)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T09139  
R:Rito, N.; Tomizawa, K.; Tanaka, K.; Matsui, M.; Kendrick, R.; Sato, T.; Nakagawa, H.  
Plant Mol. Biol. 34, 307-316, 1997  
A:Title: Characterization of 26S proteasome alfa- and beta-type and ATPase subunits from  
A:Reference number: Z16562; MUID:97351564; PMID:9207846  
A:Accession: T09139  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-249 <ITO>  
A:Cross-references: UNIPROT:O24362; UNIPARC:UPI0000132450; EMBL:D78173; NID:G2285801; PFI  
C:Superfamily: multicatalytic endopeptidase complex chain C9

Alignment Scores:  
Pred. No.: 6.39 Length: 249  
Score: 9.00 Matches: 9  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.9% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x T09139 (1-249)

QY 17 GAGGCCAAGCGCTGCTACTGCCGCC 43  
|||||  
Db 234 GUALALYSALAAAlaAlaThrAlaAla 442

RESULT 2

H84168  
molybdenum cofactor biosynthesis protein [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-feb-2001 #text\_change 09-Jul-2004  
C:Accession: H84168  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabbid  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: H84168  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-297 <STO>  
A:Cross-references: UNIPROT:Q9HST6; UNIPARC:UPI0000063541; GB:AE004437; NID:g10579734; F  
C:Genetics:  
A:Gene: moaE

Alignment Scores:  
Pred. No.: 6.22 Length: 297  
Score: 9.00 Matches: 9  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.9% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x H84168 (1-297)

Qy 48 GCAGCGCGCAGTAGCAGCGCCTTG 22  
Db 50 AlaAlaAlaValAlaAlaAlaLeu 58

RESULT 3  
S63587  
gene pacC protein - Aspergillus niger  
C:Species: Aspergillus niger  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S63587  
R:MacCabe, A.P.; van den Hombergh, J.P.T.W.; Tilburn, J.; Arst Jr., H.N.; Visser, J.  
Mol. Gen. Genet. 250, 367-374, 1996  
A:Title: Identification, cloning and analysis of the Aspergillus niger gene pacC, a wide  
A:Reference number: S63587; MUID:96180654; PMID:8602152  
A:Accession: S63587  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-667 <MAC>  
A:Cross-references: UNIPROT:Q00203; UNIPARC:UPI000006C82D; GB:X98417; GB:S81349; NID:g14  
C:Genetics:  
A:Gene: pacC

Alignment Scores:  
Pred. No.: 5.47 Length: 667  
Score: 9.00 Matches: 9  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.9% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x S63587 (1-667)

Qy 26 GCGCGTGTACTGCGCGCCTTCT 52  
Db 43 AlaAlaAlaThrAlaAlaAlaLeu 51

RESULT 4  
F70649  
probable Acyl-CoA Dehydrogenase - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: F70649  
R:Cole, S.T.; Broech, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
i Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: F70649  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-721 <COL>  
A:Cross-references: UNIPROT:P95097; UNIPARC:UPI00000D600B; GB:Z83866; GB:AL123456; NID:g1  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: fadE22

Alignment Scores:  
Pred. No.: 5.41 Length: 721  
Score: 9.00 Matches: 9  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.9% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x F70649 (1-721)

Qy 45 GCGCGCAGTAGCAGCGCCTTGCC 19  
Db 284 AlaAlaAlaValAlaAlaAlaLeu 292

RESULT 5  
BWASBE  
bimE protein - Emericella nidulans  
C:Species: Emericella nidulans, Aspergillus nidulans  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C:Accession: A37879  
R:Engle, D.B.; Osmari, S.A.; Osmari, A.H.; Roeborough, S.; Xiang, X.; Morris, N.R.  
J. Biol. Chem. 265, 16132-16137, 1990  
A:Title: A negative regulator of mitosis in Aspergillus is a putative membrane-spanning  
A:Reference number: A37879; MUID:90375468; PMID:1697851  
A:Accession: A37879  
A:Molecule type: mRNA  
A:Residues: 1-2073 <ENG>  
A:Cross-references: UNIPROT:P24686; UNIPARC:UPI0000126941; GB:M59705; GB:J05607; NID:g16  
A:Note: In addition to three predicted transmembrane domains, there are several potential  
asparaginase kinase, and one sequence that resembles a nuclear localization signal  
C:Comment: This protein is part of a regulatory pathway that includes the nimA protein k  
ter mitosis and prevent them from leaving mitosis.  
C:Genetics:  
A:Gene: bimE  
C:Superfamily: bimE protein  
C:Keywords: cell cycle control; mitosis; transmembrane protein  
F:1623-1643/Domain: transmembrane #status predicted <TM1>  
F:1685-1703/Domain: transmembrane #status predicted <TM2>  
F:1746-1764/Domain: transmembrane #status predicted <TM3>

Alignment Scores:  
Pred. No.: 4.58 Length: 2073  
Score: 9.00 Matches: 9  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.9% Indels: 0  
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x BWASBE (1-2073)

Qy 754 ATCCTGTGGCCCTCCATCTTTTACGT 728  
Db 960 IleLeuValAlaLeuHisLeuLeuArg 968

RESULT 6  
GNVEY7  
genome polyprotein - human enterovirus 70 (strain J670/71)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core prot  
NA-directed RNA polymerase (NC 2.7.7.48)  
C:Species: human enterovirus 70  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 09-Jul-2004  
C:Accession: A36253  
R:Ryan, M.D.; Jenkins, O.; Hughes, P.J.; Brown, A.; Knowles, N.J.; Booth, D.; Minor, P.D.  
J. Gen. Virol. 71, 2291-2299, 1990



A:Title: The complete nucleotide sequence of enterovirus type 70: relationships with oth  
A:Reference number: A36253; MUID:91037960; PMID:2172447

A:Accession: A36253  
A:Molecule type: genomic RNA  
A:Residues: 1-2194 <RYA>  
A:Cross-references: UNIPROT:P32537; UNIPARC:UPI00001749F5; GB:D00820  
C:Superfamily: poliovirus genome polyprotein  
C:Keywords: coat protein; core protein; genome-linked protein; hydrolase; nucleotidyltra  
F:1-69/Product: coat protein 1A #status predicted <VP4>  
F:70-319/Product: coat protein 1B #status predicted <VP2>  
F:320-561/Product: coat protein 1C #status predicted <VP3>  
F:562-871/Product: coat protein 1D #status predicted <VP1>  
F:872-1014/Product: core protein 2A #status predicted <P2A>  
F:1015-1113/Product: core protein 2B #status predicted <P2B>  
F:1114-1443/Product: core protein 2C #status predicted <P2C>  
F:1444-1532/Product: protein 3A #status predicted <P3A>  
F:1533-1554/Product: genome-linked protein VPg #status predicted <VPg>  
F:1555-1737/Product: proteinase 3C #status predicted <P3C>  
F:1738-2194/Product: RNA-directed RNA polymerase #status predicted <RRP>  
F:1535/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Alignment Scores:  
Pred. No.: 4.54 Length: 2194  
Score: 9.00 Matches: 9  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.9% Indels: 0  
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x GNNYE7 (1-2194)

QY 1022 GTCCAAAGCTGAACACTCCAGTG 1048

DB 1160 ValGlnLysLeuYsGlnLeuProVal 1168

RESULT 7

AH2228

hypothetical protein asl3383 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: AH2228  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, S.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH2228  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-85 <KUR>  
A:Cross-references: UNIPROT:Q8VRR0; UNIPARC:UPI000000CE789; GB:BA000019; PIDN:BAB75082.1;  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: asl3383

Alignment Scores:  
Pred. No.: 85.7 Length: 65  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x AH2228 (1-65)

QY 2866 AAAATTATATCACTTTAGACGAG 2843

DB 37 LysIleIleSerLeuLeuGln 44

RESULT 8

FDL4W

antifreeze protein 4 precursor - winter flounder

C:Species: Pseudopleuronectes americanus (winter flounder)  
C:Date: 01-Sep-1981 #sequence\_revision 01-Sep-1981 #text\_change 09-Jul-2004  
C:Accession: A03193  
R:Lin, Y.; Gross, J.K.  
Proc. Natl. Acad. Sci. U.S.A. 78, 2825-2829, 1981  
A:Title: Molecular cloning and characterization of winter flounder antifreeze cDNA.  
A:Reference number: A03193; MUID:81247379; PMID:6265915  
A:Accession: A03193  
A:Molecule type: mRNA  
A:Residues: 1-85 <LIN>  
A:Cross-references: UNIPROT:P02734; UNIPARC:UPI0000174278  
C:Superfamily: antifreeze protein  
C:Keywords: antifreeze  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-85/Product: antifreeze protein 4 #status predicted <MAT>

Alignment Scores:  
Pred. No.: 82.2 Length: 85  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x FDL4W (1-85)

QY 26 GCCGCTGCTACTGCCGCGCTGCT 49

DB 75 AlaAlaAlaThrAlaAlaAlaAla 82

RESULT 9

AZ2592

antifreeze protein IIA7 precursor - winter flounder  
C:Species: Pseudopleuronectes americanus (winter flounder)  
C:Date: 05-Oct-1988 #sequence\_revision 05-Oct-1988 #text\_change 09-Jul-2004  
C:Accession: A22592; S12604  
R:Gourlie, B.; Lin, Y.; Price, J.; De Vries, A.L.; Powers, D.; Huang, R.C.C.  
J. Biol. Chem. 259, 14960-14965, 1984  
A:Title: Winter flounder antifreeze proteins: a multigene family.  
A:Reference number: A22592; MUID:85054993; PMID:6548752  
A:Accession: A22592  
A:Molecule type: mRNA  
A:Residues: 1-91 <GOU>  
A:Cross-references: UNIPROT:P07835; UNIPARC:UPI0000125B4B; GB:M10148; NID:G213579; PIDN:Z  
R:Gauthier, S.; Wu, Y.; Davies, P.L.  
Nucleic Acids Res. 18, 5303, 1990  
A:Title: Nucleotide sequence of a variant antifreeze protein gene.  
A:Reference number: S12604; MUID:90384854; PMID:2402466  
A:Accession: S12604  
A:Molecule type: DNA  
A:Residues: 1-63, 'V', 65-91 <GAU>  
A:Cross-references: UNIPARC:UPI0000001382; EMBL:X53718; NID:G64213; PIDN:CAA37754.1; PID  
C:Genetics:  
A:Introns: 19/2  
C:Superfamily: antifreeze protein  
C:Keywords: antifreeze; tandem repeat  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-91/Product: antifreeze protein IIA7 #status predicted <MAT>

Alignment Scores:  
Pred. No.: 81.3 Length: 91  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x AZ2592 (1-91)

QY 26 GCCGCTGCTACTGCCGCGCTGCT 49

DB 81 AlaAlaAlaThrAlaAlaAlaAla 88

```
RESULT 10
S02376
antifreeze protein precursor - yellowtail flounder
C:Species: limanda ferruginea (yellowtail flounder)
C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S02376
R:Scott, G.K.; Davies, P.L.; Shears, M.A.; Fletcher, G.L.
Eur. J. Biochem. 168, 629-633, 1987
A:Title: Structural variations in the alanine-rich antifreeze proteins of the Pleuronectes
A:Reference number: S02376; MUID:88029483; PMID:3665937
A:Accession: S02376
A:Molecule type: mRNA
A:Residues: 1-97 <CO>
A:Cross-references: UNIPROT:P09031; UNIPARC:UPI0000125B4D; EMBL:X06356; NID:G64041; PION
F:14-48/Domain: propeptide #status predicted <PRO>
A:Note: part of this sequence, including the amino end of the mature protein, was confir
C:Superfamily: antifreeze protein
C:Keywords: antifreeze
F:1123/Domain: signal sequence #status predicted <SIG>
F:14-48/Domain: propeptide #status predicted <PRO>
F:149-96/Product: antifreeze protein #status predicted <MAT>
Alignment Scores:
Pred. No.: 80.5 Length: 97
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x S02376 (1-97)
QY 26 GCCGCTGCTACTGCCGCCGCTGCT 49
Db 58 AlaAlaAlaThrAlaAlaAlaAla 65
RESULT 11
B75632
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: B75632
R:White, O.; Elsen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75632
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <WHI>
A:Cross-references: UNIPROT:Q9RZK0; UNIPARC:UPI00000C15DC; GB:AE001826; NID:G6460827; PI
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRB0127
A:Map position: megaplasmid
A:Genome: plasmid
A:Note: plasmid MP1
C:Superfamily: Deinococcus radiodurans hypothetical protein DRB0127
Alignment Scores:
Pred. No.: 78.6 Length: 113
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x B75632 (1-113)
QY 1367 CGAGTCTGGGAGCGATCTAGTAGG 1344
Db 17 ArgValTrpGluArgSerArg 24
```

```
RESULT 12
F64083
ribosomal protein L12 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: F64083
R:Gieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gotsyne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: F64083
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-123 <TIGR>
A:Cross-references: UNIPROT:P44348; UNIPARC:UPI00001685BA; GB:U32747; GB:L42023; NID:G15
C:Superfamily: Escherichia coli ribosomal protein L12
C:Keywords: protein biosynthesis; ribosome
Alignment Scores:
Pred. No.: 77.5 Length: 123
Score: 8.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 0.8% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x F64083 (1-123)
QY 48 GCAGCGGGCGAGTAGCGGGGCC 25
Db 35 AlaAlaAlaValAlaAlaAla 42
RESULT 13
SFGP
secretin precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 24-Apr-1994 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C:Accession: B35094; A01544; A36052
R:Kopin, A.S.; Wheeler, M.B.; Leiter, A.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 2299-2303, 1990
A:Title: Secretin: structure of the precursor and tissue distribution of the mRNA.
A:Reference number: A35094; MUID:90192795; PMID:2315322
A:Accession: B35094
A:Molecule type: mRNA
A:Residues: 1-131 <KOP>
A:Cross-references: UNIPROT:P01279; UNIPARC:UPI00001357C8; GB:M31496; NID:G1646670; PIDN:
R:Mutt, V.; Jorpes, J.E.; Magnusson, S.
Eur. J. Biochem. 15, 513-519, 1970
A:Title: Structure of porcine secretin. The amino acid sequence.
A:Reference number: A91147; MUID:70282334; PMID:5465996
A:Accession: A01544
A:Molecule type: protein
A:Residues: 30-56 <MUT>
A:Cross-references: UNIPARC:UPI000002BEED
A:Note: tryptic peptides were sequenced
R:Gafvelin, G.; Joernvall, H.; Mutt, V.
Proc. Natl. Acad. Sci. U.S.A. 87, 6781-6785, 1990
A:Title: Processing of prosecretin: isolation of a secretin precursor from porcine intest
A:Reference number: A36052; MUID:90370867; PMID:2395872
A:Accession: A36052
A:Status: preliminary
A:Molecule type: protein
A:Residues: 30-59, R', 92-131 <GAF>
A:Cross-references: UNIPARC:UPI000017350D
R:Bodanzky, M.; Ondetti, M.A.; Levine, S.D.; Narayanan, V.L.; Saltza, M.V.; Sheehan, J.T.
Chem. Ind. 1966, 1757-1758, 1966
A:Title: Synthesis of a heptacosapeptide amide with the hormonal activity of secretin.
A:Reference number: A90916
A:Contents: annotation
```

A;Note: synthesis confirmed the proposed structure of the natural hormone

C;Superfamily: glucagon  
C;Keywords: amidated carboxyl end; duodenal mucosa; duplication; hormone; secretagogue  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;30-56/Product: secretin #status experimental <MAT>  
F;56/Modified site: amidated carboxyl end (Val) (amide in mature form from following gly

Alignment Scores:  
Pred. No.: 76.8 Length: 131  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x SEPG (1-131)

QY 30 CTGCTACTGCGCGCTGCTTCTT 53  
Db 7 LeuLeuLeuProLeuLeuLeu 14

#### RESULT 14

F38164  
hypothetical prtein 6 (cobo 3' region) - Pseudomonas sp.

C;Species: Pseudomonas sp.

C;Date: 31-Jul-1992 #sequence\_revision 31-Jul-1992 #text\_change 09-Dec-2002

C;Accession: F38164

R;Crouzet, J.; Levy-Schil, S.; Cameron, B.; Cauchois, L.; Rigault, S.; Rouyez, M.C.; Bla  
J. Bacteriol. 173, 6074-6087, 1991

A;Title: Nucleotide sequence and genetic analysis of a 13.1-kilobase-pair Pseudomonas de  
adenosyltransferase, cobyric acid synthase, and bifunctional cobinamide kinase-cobinami

A;Reference number: A38164; MUID:92011366; PMID:1655697

A;Accession: F38164

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-141 <CRO>

A;Cross-references: UNIPARC:UPI000013A59A; GB:M62866; NID:g151150; PIDN:AAA25782.1; PID:

C;Superfamily: Transporter DME family

#### Alignment Scores:

Pred. No.: 75.9 Length: 141  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x F38164 (1-141)

QY 48 CTCTTAGTCCGCGTTTCGCCGC 71  
Db 10 LeuLeuSerAlaAlaPheAlaAla 17

#### RESULT 15

AE2962

conserved hypothetical protein Atu3299 [imported] - Agrobacterium tumefaciens (strain C5

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004

C;Accession: AE2962

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AE2962

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-145 <KUR>

A;Cross-references: UNIPROT:Q8UAS1; UNIPARC:UPI00001647DE; GB:AE008689; PIDN:AAL44115.1;

A;Experimental source: strain C58 (Dupont)

#### C;Genetics:

A;Gene: Atu3299

A;Map position: linear chromosome

C;Superfamily: Transporter DME family

Alignment Scores:  
Pred. No.: 75.5 Length: 145  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x AE2962 (1-145)

QY 48 CTCTTAGTCCGCGTTTCGCCGC 71  
Db 13 LeuLeuSerAlaAlaPheAlaAla 20

Search completed: May 2, 2006, 03:51:57  
Job time : 122 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 03:11:14 ; Search time 137.5 Seconds  
(without alignments)  
3201.816 Million cell updates/sec

Title: US-10-825-632-2  
Perfect score: 1030  
Sequence: 1 agtgctaaagctccgagg.....agaattactaaaaaaa 3120

Scoring table

Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 60.0, Fgapext 7.0  
Delop 60.0, Delext 60.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332514

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DSV=xlp  
-DB=abss/ABSSWEB.spool/US10825632/runat\_01052006\_110235\_4423/app\_query.fasta\_1  
-DB=UniProt -QFMT=fastan -SUFFIX=oligo.n2p.rup -MINWATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bases -START=1 -END=1 -MATRIX=oligo -TRANS=human40.coi  
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p  
-USER=US10825632 @CNC\_1\_1\_466 @runat\_01052006\_110235\_4423 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DSV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=60 -XGAPEXT=60 -XGAPEXT=60 -XGAPEXT=60  
-YGAPEXT=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database : UniProt 05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	886	86.0	898	1	DPP8_HUMAN
2	186	18.1	892	1	DPP8_MOUSE
3	13	1.3	432	2	O75868_HUMAN
4	13	1.3	557	2	O5TXJ2_ANOGA
5	13	1.3	621	2	O7PTT8_ANOGA
6	13	1.3	740	2	O5TTK8_ANOGA
7	13	1.3	847	2	O6GR22_XENLA
8	13	1.3	862	1	DPP9_MOUSE
9	13	1.3	863	1	DPP9_HUMAN
10	13	1.3	886	2	O7QBK1_ANOGA
11	13	1.3	923	2	O4SBM6_TETNG
12	13	1.3	1053	2	O9VC20_DROME
13	13	1.3	1113	2	O9VC19_DROME
14	11	1.1	508	2	O75273_HUMAN
15	11	1.1	850	2	O6H9E3_9TRYP
16	11	1.1	852	2	O4Q1H9_LEIMA

17	11	1.1	1371	2	Q4P650_USFMA
18	10	1.0	289	2	Q4ZMT2_PSESY
19	10	1.0	478	2	Q9SP70_CARMA
20	10	1.0	564	2	Q8K4M2_RAT
21	10	1.0	893	2	Q9VK71_DROME
22	10	1.0	893	2	Q6NR25_DROME
23	9	0.9	87	2	Q8QFV4_LAMEL
24	9	0.9	103	2	Q9LJP7_ARATH
25	9	0.9	113	2	Q911W1_9ENTO
26	9	0.9	147	2	Q6D343_ERWCT
27	9	0.9	150	2	Q61CF4_CAEBR
28	9	0.9	154	1	PSBU_CVACA
29	9	0.9	155	2	Q5Z5HS_ORYSA
30	9	0.9	163	2	Q5Z4H9_ORYSA
31	9	0.9	172	2	Q8KAJ7_ORYSA
32	9	0.9	222	2	Q7N2V9_PHOEL
33	9	0.9	230	2	Q5BW05_SCHJA
34	9	0.9	237	2	Q852K8_ORYSA
35	9	0.9	241	2	Q7NLG5_GLOVI
36	9	0.9	248	2	Q9PWL3_PFMMA
37	9	0.9	249	1	PSA3_SFIOH
38	9	0.9	250	2	Q5YLZ5_ORYSA
39	9	0.9	256	2	Q4NVI6_9DEL
40	9	0.9	262	2	Q941I5_SORBI
41	9	0.9	265	2	Q8PHG7_XANAC
42	9	0.9	274	2	Q7P202_CHRVO
43	9	0.9	285	2	Q6N7X5_RHOPA
44	9	0.9	295	2	Q5EMP8_SORBI
45	9	0.9	297	2	Q9HST6_HALSA

ALIGNMENTS

RESULT 1  
DPP8\_HUMAN STANDARD; PRT: 898 AA.  
ID DP8\_HUMAN Q7ZAC8; Q7ZAD3; Q7ZAE1; Q81WG7; Q8NEM5; Q96JX1; Q9HBM2;  
AC Q6VLX1; Q7ZAC8; Q7ZAD3; Q7ZAE1; Q81WG7; Q8NEM5; Q96JX1; Q9HBM2;  
AC Q9HBM2; Q9HBM4; Q9HBM5; Q9NKF4;  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Dipeptidyl peptidase 8 (SC 3.4.14.5) (Dipeptidyl peptidase VIII) (DP8)  
DE (Prol) dipeptidase DPP8 (Dipeptidyl peptidase IV-related protein 1)  
DE (DPP8-1).  
GN Name=DPP8; Synonyms=DPRP1; ORFNames=MSTP097, MSTP135, MSTP141;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF 334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION, CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION, AND SUBCELLULAR LOCATION.  
RC TISSUE=Placenta;  
RX MEDLINE=20467194; PubMed=11012666;  
RA Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W., Gorrell M.D.;  
RT "Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog; DPP8.";  
RL Eur. J. Biochem. 267:6140-6150(2000).  
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME REGULATION, BIOPHYSICOCHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
RC TISSUE=Testis;  
RX PubMed=12862155; DOI=10.1042/BJ20021944;  
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akinsanya K.O.;  
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member of an emerging subgroup of serine proteases.";  
RL Biochem. J. 373:179-185(2003).





CC use as long as its content is in no way modified and this statement is not  
CC removed.

DR EMBL; AF221634; AAG29766.1; -; mRNA.  
DR EMBL; AF221635; AAG29767.1; -; mRNA.  
DR EMBL; AF221636; AAG29768.1; -; mRNA.  
DR EMBL; AF221637; AAG29769.1; -; mRNA.  
DR EMBL; AY172659; AA017261.1; -; mRNA.  
DR EMBL; AY354202; AAQ33887.1; -; mRNA.  
DR EMBL; AK000290; BAA91059.1; ALT\_FRAME; mRNA.  
DR EMBL; AK027826; BAB55395.1; ALT\_INIT; mRNA.  
DR EMBL; BC030688; AAH30688.3; -; mRNA.  
DR EMBL; BC040203; AAH40203.1; ALT\_INIT; mRNA.  
DR EMBL; AF176779; AAQ13657.1; ALT\_INIT; mRNA.  
DR EMBL; AF175225; AAQ13650.1; ALT\_FRAME; mRNA.  
DR EMBL; AF173382; AAQ13623.1; ALT\_FRAME; mRNA.

## Alignment Score:

Pred. No.:	0	Length:	898
Score:	886.00	Matches:	886
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	86.0%	Indels:	0
DB:	1	Gaps:	0

US-10-825-632-2 (1-3120) x DPP8\_HUMAN (1-898)

QY	202	GGAAATGCAATGGCGAGCAGCAATGGAAACAGACAGCTGGGTTCAGATATTGAA	261
DB	13	GlyLysCysAsnMetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGlu	32
QY	262	ACTGGCGACTGTGAGGAGAATATTGAATCACAGATCGCGCTAAATTTGAGCCTTTTAT	321
DB	33	ThrAlaAspCysGluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyr	52
QY	322	GTTCAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTGCCGATACAGAAAAATATCAT	381
DB	53	ValGluArgTyrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHis	72
QY	382	GGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTGTGAAGGAATGATCCAGAT	441
DB	73	GlyTyrMetMetAlaLysAlaProHisAspPheValLysArgAsnAspProAsp	92
QY	442	GGACCTCATTCAGACAGAACTATTACCTTCGCATGTCGTGGTGAAGACAGAGAAATACA	501
DB	93	GlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThr	112
QY	502	CTGTTTATTCGAAATCCMAAATCATCAATAGACGACGATCTTAATGCTCTCTGG	561
DB	113	LeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrp	132
QY	562	AAGCCTCTTTGGATCTTTTCAGGCAACCTCGACTATGGAATGATTCTCGAAGAA	621
DB	133	LysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGlu	152
QY	622	GAACTATTAAAGAGAAAGAAACGCATTTGGAACAGTCGGAATTCCTTACCATATCAC	681
DB	153	GluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHis	172
QY	682	CAAGGAATGGAACATTCTCTGTTTCAAGCCGGTAGTGAATTTATACGTAAGAATGGA	741
DB	173	GlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGly	192
QY	742	GGGCCAAGGATTTAGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCC	801
DB	193	GlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysPro	212
QY	802	AACATACCGATGGATCCAAATTTATGCCCCGCTGATCCAGACTGGAATTCCTTTATACAT	861
DB	213	AsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHis	232
QY	862	AGCAACGATATTCGATATCTAAATCGTAAACAGAGAAAGAGAGACTCACTTATGTG	921

DB	233	SerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrVal	252
QY	922	CACATGAGCTAGCAACATCGAAGAGATGCCAGATCAGCTGGAGTCCTACCTTGT	981
DB	253	HisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheVal	272
QY	982	CTCCAAGAAGAAATTTGATAGATATTCTGGCTATTGGTGTGTCAAAAGCTGAACAAC	1041
DB	273	LeuGlnGluGluPheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThr	292
QY	1042	CCAGTGTGTGTAATAATCTTAGAATTTCTATATGAAGAAAAATGATGAATCTGAGGTG	1101
DB	293	ProSerGlyGlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGlu	312
QY	1102	ATTATTCAATGTTATCATCCCTATGTTGGAAACAAGAGAGGCGAGATTCATTCGTTAT	1161
DB	313	IleIleHisValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrPro	332
QY	1162	AAACAGGTACAGCAAAATCTTAAGTCACCTTTTAAGATGTCAGAAATATGATGCT	1221
DB	333	LysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAla	352
QY	1222	GAAGGAGGATCATAGATGTCATAGATAAGGAACTAATCAACCTTTTCAGATCTCATTT	1281
DB	353	GluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPhe	372
QY	1282	GAAGGAGTGAATATATTGCGAGAGCTGGATGACTCTCTGAGGAAAAATATGCTGGTCC	1341
DB	373	GluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSer	392
QY	1342	ATCCTTACATAGATCGCTCCCACTACAGATAGTGTGATCTCACCTGAATATTTT	1401
DB	393	IleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPhe	412
QY	1402	ATCCAGTAGAAGATGATGTTATGGAAGCGACAGACATCATTTGAGTCAGTGCCTGATCT	1461
DB	413	IleProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSer	432
QY	1462	GTGAGCCACATTAATTTCTATGAAGAAACAACAGACATCTGGATAAATATCATCATC	1521
DB	433	ValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIle	452
QY	1522	TTTTCATGTTTTTCCCAAGTCACGAGAGGAATAGAGTTTATTTTTCCTCTGAATGC	1581
DB	453	PheHisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCys	472
QY	1582	AAACACAGTTCCTTCATTTATACAAATTTACATCTATTATTAAGGAAAGCAAAATATA	1641
DB	473	LysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLys	492
QY	1642	CGATCCAGTGTGGGCTCGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATA	1701
DB	493	ArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIle	512
QY	1702	GCAATTACATGCTGGAATGGGAAGTTCCTTGGCGCGGATGGATCTTAATATCAAGTTGAT	1761
DB	513	AlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAsp	532
QY	1762	GNACTCAGAGCGCTGATATTTTGAAGACACCAAGACTCCCTTTAGAGCATCACCTG	1821
DB	533	GluValArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeu	552
QY	1822	TACGTAGTCAGTAAATCTCTGGAGAGGTCACAAAGGCTGACTCACCTGGCTACTCA	1881
DB	553	TyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSer	572
QY	1882	CATTCTCTCTCATCAGTCAGCTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAG	1941
DB	573	HisSerCysLysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLys	592
QY	1942	AATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCAAGATGACCACTTGCAA	2001
DB	593	AsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys	612

QY 2002 ACRAAGGAATTTGGGCGACCATTTTGGATTACGAGGTCCTCTCTGACTATACCTCT 2061  
DB |||||||  
QY 613 ThrLysGluPheTrpAlaThrLeuAAspSerAlaGlyProLeuProAspGlyrThrPro 632  
QY 2062 CCAGAAATTTCTCTTTTGAAGTACTACTCGGATTTACATTTGTATGGGATGCTCTACAAG 2121  
DB |||||||  
QY 633 ProGluLeuPheSerPheGluSerThrThrGlyPheThrLeuTyThrGlyMetLeuTyTrLys 652  
QY 2122 CCTCATGATCTACAGCTCGAAGAGAAATATCTCTACTGTCTCTCTCATATATGGTGGTCTCT 2181  
DB |||||||  
QY 653 ProHisAspLeuGlnProGlyLysLysTyProThrValLeuPheIleTyThrGlyGlyPro 672  
QY 2182 CAGGTGCAATTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGGTTTCAATACCTTA 2241  
DB |||||||  
QY 673 GlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyPheArgLeuAsnThrLeu 692  
QY 2242 GCCTCTCTAGGTATGTGGTTGTAGTGATAGACACAGGGGATCTCTGCACCGAGGGCTT 2301  
DB |||||||  
QY 693 AlaSerLeuGlyTyValValValValValLeuAspAsnArgGlySerCysHisArgGlyLeu 712  
QY 2302 AAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAA 2361  
DB |||||||  
QY 713 LysPheGluGlyAlaPheLysTyLysMetGlyGlnIleGluLeuAspAspGlnValGlu 732  
QY 2362 GGACTCCAAATATCTAGCTCTCCGATATGATTTTCATGACTTGTAGCTGTGGGATCCAC 2421  
DB |||||||  
QY 733 GlyLeuGlnTyLeuAlaSerArgTyAspPheIleAspPheAspArgValGlyIleHis 752  
QY 2422 GCCTGGTCTTATGAGGATACCTCTCCGATGGCATTTATCGAGATCGATATCTTC 2481  
DB |||||||  
QY 753 GlyTrpSerTyGlyGlyTyLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePhe 772  
QY 2482 AGGTTGCTATGTCTGGGGCCCGACGTCACCTGTGTGATCTTCTATGATACAGGATACAG 2541  
DB |||||||  
QY 773 ArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyAspThrGlyTyThr 792  
QY 2542 GAACCTTATATGGTTCACCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCC 2601  
DB |||||||  
QY 793 GluArgTyMetGlyHisProAspGlnAsnGluGlnGlyTyTyLeuGlySerValAla 812  
QY 2602 ATCAAGCAGAAAGTTCCCTCTCGAACCAATCGTTTACTGCTCTTACATGGTTCTCTG 2661  
DB |||||||  
QY 813 MetGlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeu 832  
QY 2662 GATGAGATGTCCATTTTGACATACCATATATCTAGTGGTCTTTTGTAGTGGGCTGGA 2721  
DB |||||||  
QY 833 AspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGly 852  
QY 2722 AAGCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATAAAGTTCCTGAATCG 2781  
DB |||||||  
QY 853 LysProTyAspLeuGlnIleTyProGlnGluArgHisSerIleArgValProGluSer 872  
QY 2782 GGAGAACATTTAGACTGCATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGTATT 2841  
DB |||||||  
QY 873 GlyGluHisTyGluLeuHisLeuLeuHisTyLeuGlnGluAsnLeuGlySerArgIle 892  
QY 2842 GCTGCTCTAAAGTGATA 2859  
DB |||||||  
QY 893 AlaAlaLeuLysValIle 898

## RESULT 2

DPP8\_MOUSE STANDARD; PRT; 892 AA.  
AC Q80YA7; Q80Y46;  
DT 13-SEP-2005 (Rel. 48, Created)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII)  
DE (D88).  
GN Name=Dpp8;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
Murioidea; Muridae; Murinae; Mus.  
NCBI\_taxID=10090;  
[1]  
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawai H., Kawagaya Y., Kedzierzki R.M., King B.L.,  
RA Kongsava A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numa K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Sample C.A., Setou M., Shinada K.,  
RA Sultana R., Takemata Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Valardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wyshaw-Portis A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573 (2002).  
[2]  
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC STRAIN=C57BL/6J; TISSUE=Brain;  
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.A., McEwan P.J., McKernan K.J., Abramson R.D., Mullaly S.J.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Whiting M., Madan A., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smalium D.E.,  
RA Scherch A., Schein J.S., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[3]  
PHOSPHORYLATION SITE TYR-325,  
RX PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;  
RA Shu H., Chen S., Bi Q., Mumby M., Breken D.L.;  
RT "Identification of phosphoproteins and their phosphorylation sites in  
RT the WEHI-231 B lymphoma cell line."  
RL Mol. Cell. Proteomics 3:279-286 (2004).  
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal  
CC dipeptides from proteins having a pro or Ala residue at position  
CC 2. May play a role in T-cell activation and immune function (By  
CC similarity).  
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-  
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided

Zaa is neither Pro nor hydroxyproline.  
-1- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride (AEBSF), and by di-isopropyl fluorophosphate. Specifically inhibited by isoidolone derivatives (By similarity).  
-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
-1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.  
-----  
This Swiss-Prot entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
-----  
EMBL; AK016546; BAB30295.2; -; mRNA.  
EMBL; BC043124; AAH43124.1; -; mRNA.  
EMBL; BC059222; AAH59222.1; -; mRNA.  
DR MEROPS; S09.018; -;  
DR Ensembl; ENSMUSG0000032393; Mus musculus.  
DR MGI; MGI:1921638; Dpp8.  
DR InterPro; IPR001375; Peptidase\_S9.  
DR InterPro; IPR002469; Peptidase\_S9B.  
DR InterPro; IPR000379; Ser\_estrs.  
DR Pfam; PF00930; DPPIV\_N; 1.  
DR Pfam; PF00326; Peptidase\_S9; 1.  
KW Aminopeptidase; Hydrolase; Phosphorylation; Protease; Serine protease.  
FT ACT\_SITE 749 749 Charge relay system (By similarity).  
FT ACT\_SITE 827 827 Charge relay system (By similarity).  
FT ACT\_SITE 859 859 Charge relay system (By similarity).  
FT MOD\_RES 325 325 Phosphotyrosine.  
FT CONFLICT 87 87 G -> R (in Ref. 1).  
SQ SEQUENCE 892 AA; 102186 MW; 59081CD9792E03ED CRC64;  
  
Alignment Scores:  
Pred. No.: 1.39e-184 Length: 892  
Score: 186.00 Matches: 286  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 18.1% Indels: 0  
DB: 1 Gaps: 0  
  
US-10-825-632-2 (1-3120) x DPP8\_MOUSE (1-892)  
  
QY 1999 AAAAAAGAAATTTGGGCCACCATTTGGATTGAGATTGAGAGGCTCTCTCTGACTACT 2058  
Db 606 LysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThr 625  
QY 2059 CTTCCAGAAATTTCTTTTGAAGTACTACTGGATTACATTTGATGGGATCTCTAC 2118  
Db 626 ProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyr 645  
QY 2119 AAGCCTCATGATCTACAGCCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGGTGGT 2178  
Db 646 LysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 665  
QY 2179 CTTCCAGGTGAGTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC 2238  
Db 666 ProGlnValGlnLeuValAsnAsnA-gPheLysGlyValLysTyrPheArgLeuAsnThr 685  
QY 2239 CTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGGGATCTGTCCACCGAGGG 2298  
Db 686 LeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGly 705  
QY 2299 CTTAAATTTGAAGCGCTTTAAATATAAATGGGTCAATAGAAATTCAGATCAGGTG 2358  
Db 706 LeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnVal 725  
QY 2359 GAAGGACTCCAAATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATC 2418  
Db 726 GluGlyLeuGlnTyrLeuAlaSerGlnTyrAspPheIleAspLeuAspArgValGlyIle 745  
QY 2419 CACGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTTCAGATATC 2478

Db 746 HisGlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIle 765  
QY 2479 TTCAGGTTGCTATTGCTGGGCCCCCACTCTCTGTGGATCTCTATGATACAGATAC 2538  
Db 766 PheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyr 785  
QY 2539 ACGGAACGTTATATGGTGCACCTCGACCAATGAACAGGCTATTACTTAGGATCTGTG 2598  
Db 786 ThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerVal 805  
QY 2599 GCCATCAAGCAGAAAGATGTCCTCTCGAACCAATGTTTACTGCTCTTACATGGTTTC 2658  
Db 806 AlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPhe 825  
QY 2659 CTGATGAGAATGTCCATTTTGCACATACCATGATATATCTAGCTTTTGTAGTGAGGCT 2718  
Db 826 LeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAla 845  
QY 2719 GGAAGCCATATGATTTACAGATCTATCTCAGGAGACACACAGATAAGAGTTCTCTGAA 2778  
Db 846 GlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGlu 865  
QY 2779 TCGGAGAACATTTATGAACTGATCTTTTGCACTACTCTTCAAGAAACCTTGGATCAGT 2838  
Db 866 SerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArg 885  
QY 2839 ATTGCTGCTCTAAAGTGATA 2859  
Db 886 IleAlaAlaLeuLysValIle 892  
  
RESULT 3  
ID 075868 HUMAN PRELIMINARY; PRT; 432 AA.  
AC 075868;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE R33083.1 (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,  
RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,  
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,  
RA Attix C., Andreise T., Trankheim M., Amico-Keller G., Coesfield J.,  
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,  
RA Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
RA Olsen A.S., Carraro A.V.;  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC005783; AAC62840.1; -; Genomic DNA.  
DR Ensembl; ENSG00000142002; Homo sapiens.  
DR InterPro; IPR002469; Peptidase\_S9B.  
DR Pfam; PF00930; DPPIV\_N; 1.  
FT NON\_TER 432 432  
SQ SEQUENCE 432 AA; 48595 MW; 64E2B88BE0523A7E CRC64;  
  
Alignment Scores:  
Pred. No.: 0.00255 Length: 432  
Score: 13.00 Matches: 13  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 1.3% Indels: 0  
DB: 2 Gaps: 0  
  
US-10-825-632-2 (1-3120) x 075868\_HUMAN (1-432)  
  
QY 607 TATCTCGAGAGAGAACTATTATAGAGAAAGAAACGC 645



```
DR InterPro: IPR002469; Peptidase_S9B.
DR GO: GO:0008234; F: catalytic activity; IEA.
DR GO: GO:0008236; F: serine-type peptidase activity; IEA.
DR Pfam: PF00930; DPPIV_N; 1.
DR Pfam: PF00326; Peptidase_S9; 1.
FT NON TER 1
SQ SEQUENCE 740 AA; 84151 MW; D59F131BBA095CCB CRC64;

Alignment Scores:
Pred. No.: 0.0024 Length: 740
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x Q5TTK8_ANOGA (1-740)
QY 2416 ATCCACGGTGGTCCTATGAGGATACCTCTCCCTGATG 2454
|||||
DB 603 ILEHISGLYRPSERTYRGlyGlyTyrLeuSerLeuMet 615

RESULT 7
Q6GR22_XENLA
ID Q6GR22_XENLA PRELIMINARY; PRT; 847 AA.
AC Q6GR22;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC81313 protein.
GN Name=MGC81313;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus;
OC NCBI_TaxID=8355;
RN [1]
RP TISSUE=Embryo;
RC MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmitz J., Myers R.M.,
RA Butterfield V.S.N., Krywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative.";
RL Dev. Dyn. 225:384-391 (2002).
[3]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/gvdy.10174;
RA Klein S., Gerhard D.S.;
RT Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RL

DR EMBL; BC071112; AAH71112.1; -, mRNA.
DR GO: GO:0008234; F: catalytic activity; IEA.
DR GO: GO:0008236; F: serine-type peptidase activity; IEA.
DR GO: GO:0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro: IPR001375; Peptidase_S9.
DR InterPro: IPR002469; Peptidase_S9B.
DR InterPro: IPR000379; Ser_estrs.
DR Pfam: PF00930; DPPIV_N; 1.
DR Pfam: PF00326; Peptidase_S9; 1.
SQ SEQUENCE 847 AA; 97484 MW; F515F2609B267BAF CRC64;

Alignment Scores:
Pred. No.: 0.0024 Length: 847
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x Q6GR22_XENLA (1-847)
QY 607 TATTCTCGAGAGAGAACTATTAAAGAGAAAGAAACGC 645
|||||
DB 109 TyrSerArgGluGluLeuLeuArgGluuArgLysArg 121

RESULT 8
DPP9_MOUSE
ID DPP9_MOUSE STANDARD; PRT; 862 AA.
AC Q8BVG4; Q6KAM9; Q8BWT9;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)
DE (Dipeptidyl peptidase-like protein 9) (DPLP9).
GN Name=Dpp9;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Liver, and Olfactory bulb;
RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogauchi A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Brad D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numa K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sample C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Valverde R., Wagner L., Wahlestedt C., Wang Y., Wanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
```





cytoplasmic localization and DPV-like peptidase activity.";  
 Biochim. Biophys. Acta 1679:18-28(2004).  
 [4]  
 NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
 TX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzyzinski M.I., Skalska U., Smalish D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [5]  
 NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 30-863 (ISOFORM 3).  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 272-863 (ISOFORM 2), AND  
 RP PARTIAL NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 5).  
 RC TISSUE=Glial tumor, Ovary, Spleen, and Trachea;  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,  
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,  
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,  
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,  
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,  
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,  
 RA Fujimori K., Tanai H., Kinata M., Watanabe M., Hirakawa S., Chiba Y.,  
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,  
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
 RA Nomura Y., Togiani S., Komai F., Hara R., Takeuchi K., Arita M.,  
 RA Imose N., Mueshino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
 RA Moriya S., Moniyan H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe K., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45(2004).  
 [6]  
 NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 209-863 (ISOFORM 4), AND  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 298-863 (ISOFORM 2).  
 RC TISSUE=Melanoma;  
 RG The German cDNA consortium;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal  
 CC dipeptides from proteins having a Pro or Ala residue at position  
 CC 2.  
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-  
 Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided  
 Zaa is neither Pro nor hydroxyproline.  
 CC -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-  
 CC (2-aminoethyl)benzenesulphonyl fluoride (AEBSF), and by di-  
 CC isopropylfluorophosphate.  
 CC BIOPHYSICOCHEMICAL PROPERTIES:  
 CC Kinetic parameters:  
 CC KM=161 uM for Ala-Pro-AMC;  
 CC KM=180 uM for Ala-Pro-AFC;  
 CC pH dependence:  
 CC Optimum pH is 7.5-8.5. Little activity below pH 6.5;  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; cytosol.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=5;  
 CC Name=1; Synonyms=Short;  
 CC IsoId=Q86T12-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Long;  
 CC IsoId=Q86T12-2; Sequence=VSP\_013865;  
 CC Note=Incomplete sequence;  
 CC Name=3;  
 CC IsoId=Q86T12-3; Sequence=VSP\_013867, VSP\_013868;  
 CC Name=4;  
 CC IsoId=Q86T12-4; Sequence=VSP\_013869;  
 CC Name=5;  
 CC IsoId=Q86T12-5; Sequence=VSP\_013865, VSP\_013866;  
 CC Note=Incomplete sequence;  
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in  
 CC liver, heart and muscle, and lowest levels in brain.  
 CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPVIV subfamily.  
 CC -!- CAUTION: Ref.6 (CAD39039) sequence differs from that shown due to  
 CC frameshifts in positions 432 and 460.  
 -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 -----  
 EMBL: AF452102; AAL47179.1; -; mRNA.  
 EMBL: AY126660; AA017362.1; -; mRNA.  
 EMBL: AF542510; AA073880.2; -; mRNA.  
 EMBL: AY374518; AA083119.1; -; mRNA.  
 EMBL: BC000970; AAH00970.1; -; mRNA.  
 EMBL: BC037948; AAH37948.1; -; mRNA.  
 EMBL: AK054656; BAB70784.1; ALT INIT; mRNA.  
 EMBL: AK075030; BAC11362.1; -; mRNA.  
 EMBL: AK131100; BAC85150.1; -; mRNA.  
 EMBL: AK131499; BADI8643.1; ALT INIT; mRNA.  
 EMBL: AL834376; CAD39039.3; ALT\_FRAME; mRNA.  
 EMBL: CR627380; CAH10477.1; -; mRNA.  
 MEROFS; S09.019; -;  
 DR Ensembl; ENSG00000142002; Homo sapiens.  
 DR HGNC; HGNC:18648; DPP9.  
 DR MIM; 608258; -;  
 DR InterPro; IPR001375; Peptidase\_S9.  
 DR InterPro; IPR002469; Peptidase\_S9B.  
 DR InterPro; IPR000379; Ser\_estr.  
 DR Pfam; PF00930; DPVIV\_N; 1.  
 DR Pfam; PF00326; Peptidase\_S9; 1.  
 KW Alternative splicing; Aminopeptidase; Hydrolase; Protease;  
 KW Serine protease.  
 FT ACT SITE 730 730 Charge relay system (By similarity).  
 FT ACT SITE 808 808 Charge relay system (By similarity).  
 FT ACT SITE 840 840 Charge relay system (By similarity).  
 FT VARSPLIC 1 1 M -> LSRAYPCVRCRPPPLPGLPGSQSRMRDREAPLD  
 FT PGPAGSGRPTSRVSUSHACSWNGSLDPLECTPALLRSAE  
 FT RLMKVKKLDKENTGWSRSLNSEGAER (in  
 FT isoform 2 and isoform 5).  
 FT /FTid=VSP\_013865.  
 FT Missing (in isoform 5).  
 FT /FTid=VSP\_013866.  
 FT QLVNNSFGIKYRLRLNTLASLGYAV -> SAHLLPRPPPHH  
 FT PPEDSPSLKCOL (in isoform 3).  
 FT



```
FT FT VARSPLIC 675 863 /FTid=VSP_013867.
FT FT Missing (in isoform 3).
FT FT /FTid=VSP_013868.
FT FT Missing (in isoform 4).
FT FT /FTid=VSP_013869.
FT FT I -> N (in Ref. 3).
FT FT CONFLICT 204 204
FT FT CONFLICT 571 571 C -> W (in Ref. 5); BAC85150).
FT FT CONFLICT 709 709 L -> P (in Ref. 5); BAB18643).
FT FT CONFLICT 753 753 G -> C (in Ref. 5); BAB70784).
SQ SEQUENCE 863 AA; 98263 MW; 40FE0B78E26CDED5 CRC64;

Alignment Scores:
Pred. No.: 0.0024 Length: 863
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x DPP9_HUMAN (1-863)
QY 607 TATTCTCGAGAGAACTATTAAAGAGAAAGAAACGC 645
Db 121 TyrSerArgGluGluLeuLeuArgGluArgLysArg 133

RESULT 10
Q7QBK1 ANOGA
ID Q7QBK1 ANOGA PRELIMINARY; PRT; 886 AA.
AC Q7QBK1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE ENSANGP0000016526 (Fragment).
GN ORFNames=ENSANGG0000014037;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAB01008879; EAA08416.2; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; Fcatalytic activity; IEA.
DR GO; GO:0008236; Fserine-type peptidase activity; IEA.
DR GO; GO:0006508; Pproteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estra.
DR Pfam; PF00930; DPPIV N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON TER 1
FT NON TER 886
SQ SEQUENCE 886 AA; 100440 MW; 3C284605CAA57DB4 CRC64;

Alignment Scores:
Pred. No.: 0.00239 Length: 886
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
```

```
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x Q7QBK1 ANOGA (1-886)
QY 2416 ATCCACGGCTGCTCTTATGAGGATACCTCTCCCTCATG 2454
Db 750 ILEHISGLYTRPSERTYGLYGLYTYRLEUSERLEUMET 762

RESULT 11
Q4SBM6 TETNG
ID Q4SBM6 TETNG PRELIMINARY; PRT; 923 AA.
AC Q4SBM6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 15 SCAF14667, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00020903001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dessat C., Segurens B.,
RA Anhouard V., Salanoubat M., Levy M., Baudet N., Castellano S.,
RA Bismont C., Skalli Z., Cattolico L., Katinka M., Vacherie B.,
RA Craud C., Duprat S., Brottier P., Poulain J., De Berardinis V.,
RA Pauric M., Lardier G., Chappie C., Coutanceau J.P., Guzy J., Bosak S.,
RA Lindblad-Fon K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander E., Weissenbach J., Roest Crollius H.,
RA Wincker P., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RI Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype."
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAE01014667; CAG01956.1; -; Genomic_DNA.
FT NON TER 1
FT NON TER 923
SQ SEQUENCE 923 AA; 105211 MW; 1397023B2004D009 CRC64;

Alignment Scores:
Pred. No.: 0.00239 Length: 923
Score: 13.00 Matches: 13
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 1.3% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x Q4SBM6 TETNG (1-923)
QY 607 TATTCTCGAGAGAACTATTAAAGAGAAAGAAACGC 645
Db 112 TyrSerArgGluGluLeuLeuArgGluArgLysArg 124

RESULT 12
Q9VC20 DROME
ID Q9VC20 DROME PRELIMINARY; PRT; 1053 AA.
AC Q9VC20; Q81H07;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
```

10-MAY-2005 (TRENBLrel. 30, Last annotation update)  
CG3744-PA, isoform A (CG3744-pc, isoform c) (LD33755p).  
ORFNames=CG3744;  
Drosophila melanogaster (Fruit Fly).  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_TaxID=7227;  
RN [1]  
RN NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lako P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle B.J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirska R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RN NUCLEOTIDE SEQUENCE.  
RX MEDLINE=2426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Pacle B.J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirska R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RN NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirska R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
a genomics perspective.";  
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
RN [4]  
RN NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426069; PubMed=12537572;

RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA BattenCourt B.R., Celniker S.E., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
systematic review.";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [5]  
RN NUCLEOTIDE SEQUENCE.  
RX Berkeley Drosophila Genome Project;  
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
RA Hoskins R., Stapleton M., Pacle B.J., Park S., Svirska R., Smith E.,  
RA Yu C., Rubin G.;  
RT "Drosophila melanogaster release 4 sequence.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RN NUCLEOTIDE SEQUENCE.  
RX FlyBase;  
RA Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=Berkley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacle B.J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- INTERACTION:  
CC Q9VZRO:CG12016; NbExp=1; IntAct=EBI-105926, EBI-152819;  
DR EMBL; AE003749; AAP56357.2; -; Genomic DNA.  
DR EMBL; BT001499; AAN71254.1; -; mRNA.  
DR IntAct; Q9VC20; -;  
DR MEROPS; S09.016; -;  
DR Ensembl; CG3744; Drosophila melanogaster.  
DR FlyBase; FBGN0039240; CG3744.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:004274; F:dipeptidyl-peptidase IV activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001375; Peptidase\_S9.  
DR InterPro; IPR002469; Peptidase\_S9B.  
DR InterPro; IPR000379; Ser esters.  
DR Pfam; PF00930; DPPIV\_N; I.  
DR Pfam; PF00326; Peptidase\_S9; 1.  
SQ SEQUENCE 1053 AA; 118053 MW; C94AA663AB464577 CRC64;  
  
Alignment Scores:  
Pred. No.: 0.00236 Length: 1053  
Score: 13.00 Matches: 13  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 1.3% Indels: 0  
DB: 2 Gaps: 0  
  
US-10-825-632-2 (1-3120) x Q9VC20\_DROME (1-1053)  
  
QY 2416 ATCCAGCGCTGGTCTATGAGGATACCTCTCCCTGATG 2454  
|||||  
Db 917 IIEHISGLYTPSPSTYRGLYTYRLEUSERLEUMET 929  
|||||  
RESULT 13  
Q9VC19\_DROME AC  
ID Q9VC19\_DROME PRELIMINARY; PRT; 1113 AA.  
AC Q9VC19;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE CG3744-PB, isoform B.



Mon May 8 13:41:51 2006

DR MEROPS; S09.019; -  
DR Ensembl; ENSG0000142002; Homo sapiens.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR InterPro; IPR001375; Peptidase\_S9  
DR InterPro; IPR002469; Peptidase\_S9B.  
DR InterPro; IPR000379; Ser\_estrs.  
DR Pfam; PF00930; DPPIV N; 1.  
DR Pfam; PF00326; Peptidase\_S9; 1.  
FT NON TER 1 1  
SQ SEQUENCE 508 AA; 57750 MW; 2F4D6645BE2D2C89 CRC64;

Alignment Scores:  
Pred. No.: 0.313 Length: 508  
Score: 11.00 Matches: 11  
Percent Similarity: 100.0% Conservatives: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 1.1% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x O75273\_HUMAN (1-508)  
QY 1804 CCCTTAGCATCACCTGACGTAGTCAGTTAC 1836  
Db 167 ProLeuGluHisIseuTyValValSerTyr 177

RESULT 15  
Q6H9E3\_9TRYP  
ID Q6H9E3\_9TRYP PRELIMINARY; PRT; 850 AA.  
AC Q6H9E3;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Trypanosoma brucei dipeptidyl peptidase.  
GN Name=dpptb;  
OS Trypanosoma brucei.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5691;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Bastos I.M.D., Girard D., Santana J., Grellier P.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ617830; CA92394.1; -; mRNA.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001375; Peptidase\_S9.  
DR InterPro; IPR002469; Peptidase\_S9B.  
DR InterPro; IPR000379; Ser\_estrs.  
DR Pfam; PF00930; DPPIV N; 1.  
DR Pfam; PF00326; Peptidase\_S9; 1.  
DR PRINTS; PR00453; VWFADOMAIN.  
SQ SEQUENCE 850 AA; 93842 MW; FB1D793DD9939BD4 CRC64;

Alignment Scores:  
Pred. No.: 0.299 Length: 850  
Score: 11.00 Matches: 11  
Percent Similarity: 100.0% Conservatives: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 1.1% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x Q6H9E3\_9TRYP (1-850)  
QY 2524 TATGATACAGGATACACGGAACGTTATACGGT 2556  
Db 731 TyrAspThrGlyTyrThrGluArgTyrMetGly 741

Search completed: May 2, 2006, 03:48:44  
Job time : 725.5 secs

**THIS PAGE BLANK (USPTO)**





Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGAGAATATTGAATCACAGATCGCGCTAAATTCGAGCCTTTTATGTTGAGGGTAT 333  
Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyValGluArgTyr 40  
QY 334 TCCTGGAGTCAAGCTTAAAGCTGCTTGGCCGATACAGAGAAATATATGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 CTAAGGACACACATGATTTTCATGTTGTCAGAGAGAAATGATCCAGATGGACCTCATCA 453  
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 GACAGATCTATTACCTTGCATGCTGTGTGAGACAGAGAAAAATACACTGTTTATCT 513  
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
QY 514 GAATATCCAAAATCATCAATAGACGACGAGTCTTAATGCTCTCTTGAAGCCCTTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATCTCGAGAGAAAGAACTATTAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGlnLeuLeuArg 140  
QY 634 GAAAGAAAAAGCATTTGGAACAGTCGGAATTTGCTTACGATTATCACCAAGGAGGGA 693  
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATACGTAAAGATGGAGGCCACRAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTACGCAACACCTTTAGGCCCAATCTAGTGGAACTAGTCTCCCAACATACGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCCAAAATATGCCCGCTGATCCAGACTCGATGCTTTTATACATAGCAACGATATT 873  
Db 201 AspProLysCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATCTTAACATCTGATACACAGAGAAAGAGACTCATTTATGTGCAATAGACTA 933  
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240  
QY 934 GCCAACATGGAAGAATGCCAGATCAGCTGGAGTCGCTACTTTTGTCTCCAAGAAGAA 993  
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260  
QY 994 TTTGATAGATATTCTCGCTATTGGTGTGTCCTCAAAAGCTGAAACACTCCCGAGTGGT 1053  
Db 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATCTTAGAATTTATATGAAGAAATGATGAATCTGAGTGGGAAATTTATTCATGTT 1113  
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300  
QY 1114 ACATCCCTATGTTGAAACAAAGGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTACA 1173  
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAAACTCAAGTCACTTTTAAAGATGTCAGAAATATGATGATGTCAGAGCAAGATC 1233  
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGATAAGGAATTAATCAACTTTTCAGATCTTATTTAGAGAGTTGAA 1293  
Db 341 IleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
QY 1294 TATATTGCCAGAGCTGGATGACTCTCTGAGGAGAAATATGCTGGTCCATCTACATGAT 1353  
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380

QY 1354 GCCTCCAGAGCTCCCTTACAGATAGTGTGTGATCTCACCTGAATATTATATCCAGTAGAA 1413  
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATGATGTTATGGAAAGCGCAGAGACTCATGTAGTCTAGTCCCTGATTTCTGTGACGCCACTA 1473  
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATGAAGAAAAACAACAGACATCTGGATAAATATCATGACATCTTTTCATGTTTT 1533  
Db 421 IleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisValPhe 440  
QY 1534 CCCCAGAGTCACGAAGAGGAAATTTAGTGTATTTTTCCTCTGAAATGCAGAAACAGGTTTC 1593  
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
QY 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGT 1653  
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
QY 1654 GGGCTGCCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713  
Db 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500  
QY 1714 GGTGAATGGGAAGTCTTGGCCGCGCATCGATCTAATATCCAAAGTTGATGAAGCAGAGG 1773  
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520  
QY 1774 CTGTTATTTTGAAGGACCCAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACGT 1833  
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540  
QY 1834 TAGCTAAATCTCGAGAGGTGACAAAGCTGACTGACCTGGCTACTCACATTTCTTGTGTC 1893  
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560  
QY 1894 ATCAGTCAGACACTGTCACTCTTTTATAAGTAAAGTATAGTAAACCAAGAGAAATCCACACTGT 1953  
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
QY 1954 GTGTCCCTTTTACAAAGCTATCAAGTCTCTGAAAGTACCCCACTTCGCAAAACAAGAAATTT 2013  
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
QY 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTCTGACTATCTCTCCAGAAATTTTC 2073  
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620  
QY 2074 TCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA 2133  
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
QY 2134 CAGCTGGAAAGAAATATCCCTACTGCTGTTCATATATGTTGCTCTCAGGTGCAGTTG 2193  
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660  
QY 2194 GTCAATAATCCGTTTAAAGGAGTCAAGTATTTCCTGCTGTAATACCTTAGGCTCTTAGGT 2253  
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680  
QY 2254 TATGCTGTTGATGATAGACAACAGGGGATCTCTGTCCACCGAGGCTTAAATTTGAAGGC 2313  
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700  
QY 2314 GCCTTTTAAATATAAATGGGTCAAATAGAAATTCAGATCAGGTGGAGGACTCCAATAT 2373  
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720  
QY 2374 CTAGCTCTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCAATCCACGGCTCGTCTAT 2433  
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyr 740

```
QY 2434 GGAGGATACCTCCCTGATGCATTAAATGACAGAGTCCAGATATCTTCAGGGTGTCTATT 2493
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTTATGATACAGGATACACGGAACTTATATG 2553
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCCAGCAGCAAGTAAACAGGGCTATTACTTAGAGTCTGGCCATGCAAGCAGAA 2613
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
781 GlyHisProAspGlnAsnGluGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCTCCCTCAACCAATCGTTTACTGCTCTTACATGCTTTCCCTGGATGAGAAATGTC 2673
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCACATACCAAGTATATTACTGAGTCTTTTGTAGGGCTGGAAGCCATATGAT 2733
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGGAGACACAGCATTAAGATTCCTGAAATCGGGAACATTAT 2793
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
841 LeuGlnIleTyrProGlnArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAAGTGCATCTTTTGCACCTACCTTCAAGAAACCTTTGGATCACGTTATGCTGCTCTAAAA 2853
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLeuLys 880
QY 2854 GTGATA 2859
Db |||||||
881 ValIle 882

RESULT 2
US-10-070-464-1
; Sequence 1, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-1

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-1 (1-882)
QY 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAAACTCGCGACTGT 273
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 MetAlaAlaIleMetGluThrGlnGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATGAATCACAGGATCGGCCTTAAATTCGAGCCCTTTTATGTTGAGCGGTAT 333
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGAGTCAGCTTAAAAAGCTGCTGCGATACACGAAAAATATCATGCTCATCATGATG 393
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTC 453
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTTATACCTTCCCATGCTGCTGAGAACACAGAGAAAAATACACTGTTTATTCT 513
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCCAAACTATCAATAGACGACGAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAGAAGAAGACTTATAGA 633
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAAAGCATTGGAACAGTCGGAATTCCTTACGATTATTCACCAAGGAAGTGA 693
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATACGTAAAAAGATGGAGGCCCAAGGA 753
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGCTCCCAACATACGGATG 813
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTCGATTGCTTTTATACATAGCAACGATATT 873
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCTGTAACACAGAGAAGAAAGAGACTCCTATTGTGCACAATGAGCTA 933
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
221 TrpIleSerAsnIleValThrArgGluGluArgValAlaThrPheValLeuGlnGluGlu 240
QY 934 GCCAATATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAGAA 993
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTTCATAGATATTCTGCTATTGCTGCTGCTCCAAAAGCTGAAAACAACTCCCAAGTGGTGGT 1053
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTTATGTTGGAACACAGAGGGAGATTCATTCCTCGTTATCTTAAACACAGTACA 1173
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACCTTTTAAAGATGTGAGAAATAATGATGTGCTGAAGGAAGATC 1233
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCTAGATAAGGAACCTAATTCACCTTTTGGAGATCTCTATTGGAAGGAGTTGAA 1293
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGATCGACTCTCTGAGGGAAATATGCTTGGTCCATCTCTACTAGAT 1353
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCCAAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
```

```
QY 1414 GATGATGTTATGAAGGACGAGACTCATTGAGTCAGTCGCTGATTCCTGTCGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCATGACATCTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTACAGGAAGGAAATTTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysIleThrGlyPhe 460
QY 1594 CGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATAAAGCATCCAGTGGT 1653
Db 461 ArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGCCTGCTGCTCAAGTGATTTCAAGTCTCTATCAAGAGGAGATAGCAATTTACCAAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGCCCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG 1773
Db 501 GlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGluValArg 520
QY 1774 CTGTATATTTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGluThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TAGCTAAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys 560
QY 1894 ATCAGTCAGCAGCTGTCATCTTCTTATAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTACAAGCTATCAAGTCTGAAGATGACCACTCTGCAAAACAAGGATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGATTCACAGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2073
Db 601 TipAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CACCTTGGAAAGAAATATCCTACTGTGCTGTCTATATATGCTGCTCAGGTGCAAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGCTGTTGATGATACAGAACAGGGGATCCTGTCCACCGGGGCTTAAATTTGAAGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATGGCTCAAAATAGAAATTCACCATCAGGTGGAAGACTCCAAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTCTCCGATATGATTTTCAATGACTTACATGCTGTGGGATCCACCGCTGCTCAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
QY 2434 GGAGATACCTCTCCCTGATGGCAATTAATGAGAGGTTCAGATATCTTCAGGGTTCCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
```

```
QY 2494 CTTGGGCCCCAGTCACTCTCTGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuThrPheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTCCACATACCACTATATTACTAGTCTTTTGTAGTGGCTTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCTCAGAGAGACACAGATAAGAGTTCTCTGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACCTGATCTTTTGTGCACTACCTTCAAGAAACCTTGGATCACGTTATGCTGCTTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
RESULT 3
US-09-976-674-19
; Sequence 19 Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19
Alignment Scores:
Pred. No.: 0 Length: 658
Score: 655.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.6% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x US-09-976-674-19 (1-658)
QY 214 ATGGGACGAGCAATCGAAACAGACAGCTGGCTTTCAGATATTGAACTGCGGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGAATCAGATCGGCCTTAATTTGAGCGCTTTTATGTTAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCTCTGAGTCAGCTTAAAGACCTCTTGGCGATACCGAAAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
```

```
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTC 453
Db 61 AlaiysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTCCATGCTGCTGTGAGACAGAGAAAATACACTGTTTATCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAAACATCAATAGACAGCAGTCTTAATGCTCTCTTTGAAAGCCTCTTTTG 573
Db 101 GluileProLysThrIleAsnArgAlaAlaValLeuMetLeuSerIleProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACATGACTATGGAATGTATTTCTCGAGAAAGAACTATTAA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACGCATTTGGAACAGTCGGAATTTGCTTCTTACGATTAATACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCATCGTAAAGATGGAGGCCCAAGGA 753
Db 161 ThrLeuPhePheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCCAAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAACACAGAGAAAGAGACTCACATTATGTGCACAAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTTCTCCAAAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTTCTGCTATTTGGTGTGTCGAGTCCAAAGCTGAAACAACTCCAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAAATTTATATGAAAGAAATGATGATCTGAGCTGGAAATTTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACACAGGAGGAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAGTCACCTTTTAAGATGTGACAGAAATATGATTTGATCTTATTTGAAGGAGTTGAA 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTATAGATAAGGAACTAAATCAACCTTTTGAATTTCTATTTTGAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCAGAGCTGGATGGACTCTCGAGGAAAATATGCTTGTGTCATCTCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrIleLeuLeuAsp 380
QY 1354 CDTCCAGACTCCCTACAGATAGTTGATCTCACCTGGAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATCATGCTTATGGAAGGACAGACTCATTTGAGTCAGTCCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
```

```
Db 421 IleIleTyrGluGluThrThrAspIleThrPheAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCCAGAGAGGAAATTTGATTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAGATGATTTCAAGTGTCTTATCAAGAGGAGATACCAATTTACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGGCCGGCATGATCTAATATCAAGTTGATGATGATCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGTGATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCCTCTGACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACGTAAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTTCTTTATAGTAAGTAGTAACCAAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTTCAGAGCTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTTACATGTATGGGATGCTCTACAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAAGAAATATCTCTACTGTCTGCTTTCATATATGTTGGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

RESULT 4
US-09-976-674-11
; Sequence 11, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR APPLICATION NUMBER: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-11

Alignment Scores: 0 Length: 661
Pred. No.: 655 Matches: 655
Score: 655.00
```

Percent Similarity: 100.0%  
Best Local Similarity: 100.0%  
Query Match: 63.6%  
DB: 2  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-11 (1-661)

QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGGTTGAGATATTTGAAACTGCCGACTGT 273  
DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGGAATATGTAATCAGAGATCGGCTTAAATTCGAGCCCTTTTATGTTGAGCGGTAT 333  
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACCAGAAAAATATCATCGGTACATGATG 393  
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453  
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 CACAGAATCTATTACCTTGCATGCTCTGGTGAGAACAGAGAAAAATACACTGTTTATTCT 513  
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
QY 514 GAAATTCCTCAAAATCATCAATAGACGAGCTTAAATGCTCTCTTGGAGCCCTTTTG 573  
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTCAGGCACACTGGACTATGGAATGATCTTCGAGAGAGAACTATTAAAGA 633  
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
QY 634 GAAAGAAAAGCATGGAACAGCTCGAATTCCTCTTACGATATATCAACAAGGAAGTGA 693  
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTCTGTTTCAAGCCGTAGTGAAATTTATACGTAAAGATGAGAGGCCCAACAGGA 753  
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACTAGTGTGCCAATACGGAATG 813  
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATAT 873  
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTA 933  
DB 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240  
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACTTGTGTTCTCCAGAGAA 993  
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260  
QY 994 TTTTCATAGATATTTCTCGCTATTGGTGTCCAAAGCTGAAACAACTCCCGAGTGGTGT 1053  
DB 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATCTTGAATTTCTATAGNAATATGATCTGAGGTGGAAATTTATTCATGTT 1113  
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleVal 300  
QY 1114 ACATCCCTATGTTGGAACAGAGGGGAGATTCATTCCTCGTTATCTCTAAACAGGTACA 1173  
DB 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAATTCCTAAAGTCACTTTTAAAGATGTCAGAAAAATGATGATGCTGGAAGAGATC 1233

DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGAGATTCTATTGAAAGGAGTTGAA 1293  
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
QY 1294 TATATTCCAGAGCTCGATGGACTCTCTGAGGAAAAATATGCTTGGTCCACTCTACTAGAT 1353  
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380  
QY 1354 CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACCTGGAATTTATTTATCCAGTAGAA 1413  
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATGATGTTATGGAAGGACAGAGACTCATTTAGTCACTGCTGATCTGTGACGCCACTA 1473  
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATCAAGAAACACAGACATCTGATATAATATATCCATGACATCTTTTCATGTTT 1533  
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
QY 1534 CCCCAAGTCCACGAAGAGGAATTTGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC 1593  
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
QY 1594 CGTCATTTATACAAAATACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGT 1653  
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
QY 1654 GGGCTGCTCTCAAGTGATTTCAAGTGTCTCTATCAAGAGAGATAGCAATTTACAGT 1713  
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500  
QY 1714 GGTGAATGGAAATTTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAAGT 1773  
DB 501 GlyIleTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520  
QY 1774 CTGTATATTTTGAAGGCCCAAGAGCTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833  
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540  
QY 1834 TACGTAATCCTCGAGAGGTGACAAAGCTGACCGTGGCTACTCACATTTCTTGTCTGC 1893  
DB 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys 560  
QY 1894 ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAAACCAAGAAATTCACACTGT 1953  
DB 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
QY 1954 GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCAAAACAAAGAAATTT 2013  
DB 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
QY 2014 TGGGCCCACTTTTGGATTTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073  
DB 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620  
QY 2074 TCTTTTGAAGTACTACTGGATTTTACATGATGGATGCTCTACAGCTCATGATCTA 2133  
DB 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
QY 2134 CAGCTCGGAAGAAATATCTTACTGTGCTGTTTCATATATGTTGT 2178  
DB 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

RESULT 5  
US-09-976-674-7  
; Sequence 7, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen

```
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-7

Alignment Scores:
Pred. No.: 0          Length: 690
Score: 655.00        Matches: 655
Percent Similarity: 100.0%  Conservative: 0
Best Local Similarity: 100.0%  Mismatches: 0
Query Match: 63.6%      Indels: 0
DB: 2                  Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-7 (1-690)
QY 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTCAAACTCGCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGluLeuGluValGluLeuPheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACAGAAAATATATCGGTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGACACATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GNCAGAACTTATTACCTTGCCTGCTGCTGAGAACAGAGAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCACCAACTATCAATAGACAGCAGCTTAAATGCTCTCTTGGAGGCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAGAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140
QY 634 GAAAGAAACCGATTGGAAACAGTCGGAATTCCTTACGATTATCCAGAGAAAGTGGGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTCTGTTTCAAGCCGAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGCCAAACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCCCAACATCGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATATGCTCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTTAACATCGTAACAGAGAGAAAGAGACTCATCTTATGTGCACATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnGluLeu 240
```

```
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGTGGAGTCGCTACCTTTGTTCTCCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGCTGCTCCAAAGCTGCAAACTCCAGAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTCTTAGAATTTCTATATGAAAGAAATGATGATCTGAGGTGGAAATTTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTCTATGTTGGAACAAGAGGCGAGATTCATTCGGTTATCTCTAAACACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAACTCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGGAAGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAACTCAACCTTTTGGAGATTTCTATTTGAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGAGTCTCTGAGGAAATATGCTTGTGCTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTTACAGATAGTGTTCATCTCACCTGATTTATTTATTTATTTATTTATTT 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGAGGAGAGACTCATGTAGTCCCTGATCTGTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACACACATCTGTATAAATATCCATGACATCTTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCTCAGGAAGAGAAATTTGAGTTTATTTTTCCTCTGAAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIleAlaSerGluCysLysThrGlyPhe 460
QY 1594 GCTCATTTATACAAAATATCATCTATTTTTAAAGAAAGCAAAATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTGTTCNAAGTGTCTCTATCANAAGAGGAGATAGCAATTACCAAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGGCCGCTATGATCTTAATATCAAGTTGATGAAGTCAGAAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGACATCACCTGTACGTACTGCTAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACGTAAATCTCGAGAGGTCACAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGTC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAAGCTGCTGTGACTTCTTTTATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTCAAGCTTATCAAGTCTCTGAGAGATGACCCCAACTGCAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPhe 600
```



QY 2014 TGGGCCACCAATTTGGATTACAGAGTCCCTCTCTGACTACTACTCTCCACGAAATTTTC 2073  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
601 TTPAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620  
QY 2074 TCTTTTGAAGTACTACTCGGATTTACATTTGTATGGGATGCTTACAGGCTCATGATCTA 2133  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrPheProHisAspLeu 640  
QY 2134 CAGCTGGAAAGAAATATCTTACGTCTCTCTTTCATATATGGTGGT 2178  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

## RESULT 6

US-09-976-674-21  
; Sequence 21, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 613  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-21

Alignment Scores:  
Pred. No.: 0 Length: 613  
Score: 607.00 Matches: 607  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 58.9% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-21 (1-613)

QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTGCGCACTGT 273  
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGAGATATTGAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGACGGTAT 333  
Db 21 GluGluAenileGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACAGAAATATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 GCTAAGCAGCACATGATTTCAATGTTGTGACAGCAATGATCCAGATGACCTCATTC 453  
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 GACGATCTATTACCTTGGCATGCTGTGTGAGAACAGAGAAATATACACTGTTTATCT 513  
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer 100  
QY 514 GAAATTCGCAAAATCTAATAGACGACGCTTAAATGCTCTCTTGGAGCCCTCTTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTTCAGGCACACTGGACTATGGAATGTATCTCGAGAAGAAGAACTATTAAGA 633  
Db 121 AspleuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140

QY 634 GAAAGAAAACCATTTGGAAACAGTCGGAATTTGCTTCTTACGATTATCACCAGGAAGTGA 693  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTCCTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAAGGA 753  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTACGCAACCAACCTTTAAAGCCCAATCTAGTGGAAATCTAGTTGCTCCCAACATACGGATG 813  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGATGCTTTTATACATAGCAACCATATT 873  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTAACATCGTAACACAGAGAAGAAAGAGCACTCATTTATGTCGACAAATGAGCTA 933  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
221 TrpIleSerAenileValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240  
QY 934 GCCAATCTGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCTCACAAGAA 993  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluLeu 260  
QY 994 TTTGATAGATATTTCTGGCTATTTGCTGCTCCAAAGCTGAAACAACTCCAGCTGGTGGT 1053  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATCTTAGAATTTCTATATGAAGAAAATGATGATCTGAGGTGGAAATTTATTCATGTT 1113  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
281 LysIleLeuArgIleLeuTyrGluGluAenAspGluSerGluValGluIleHisVal 300  
QY 1114 ACATCCCTATGTTGGAACCAAGAGGGCAGATTCATTCCTGTTATCTCTAAACAGGTACA 1173  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGGAAGGAAGATC 1233  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGGAGATTTCTATTGGAAGGATTTGAA 1293  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
QY 1294 TATATTCAGAGCTCGATGGACTCTCGAGGGAATATGCTTGGTCCATCTACTAGAT 1353  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380  
QY 1354 CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATGATGTTATGGAAGGAGAGACTCATGAGTCAGTGCCTGATTCGTGACGCCCTA 1473  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1533  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 IleIleTyrGluIuThrThrAspIleTrpIleAenileHisAspIlePheHisValPhe 440  
QY 1534 CCCCAAAGTCAGAGAGAAATTCAGTTTATTTTTCCTCTGAATCAAAACAGGTTTC 1593  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
QY 1594 CGTCATTTATACAAAATTCATCTATTTTAAAGGAAAGCAATATAAACCATCCAGTGGT 1653  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
QY 1654 GGGCTCGCTCGTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACCAGT 1713  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500  
QY 1714 GGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1773



```
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValA-gArg 520
QY 1774 CTGATATATTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCTTGTGTC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCACTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTACAGCTATCAAGTCTGAGAGTGCACCACTTCGCAAAACAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCACTTTTGGATTCA 2034
Db 601 TrpAlaThrIleLeuAspSer 607

RESULT 7
US-10-070-464-5
; Sequence 5, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-5

Alignment Scores:
Pred. No.: 1.04e-313 Length: 465
Score: 341.00 Matches: 341
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 33.1% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-5 (1-465)
QY 1165 ACAGTACAGCAATCTCTAAAGTCATCTTTTAAGATCTCAGAAATATGATTGATCTGCA 1224
Db 1 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 20
QY 1225 GDAAGGATCATAGATGTCATAGATAAGGAATCAACCTTTTGGAGATCTATTGAA 1284
Db 21 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 40
QY 1285 GGAGTTGAATATATTGCGAGCTGGAATGACTCTCGAGGAAAAATATCTTGGTCCATC 1344
Db 41 GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle 60
QY 1345 CTACTAGATCGCTCCGAGCTCGCCTACAGATGTTGATCTCACCTGGAATTTATTC 1404
Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValIleLeuSerProGluLeuPheIle 80
```

```
QY 1405 CCAGTAGAAGATCATGTTTATGGAAGGACAGAGACTCATTTGAGTCAGTGCCTGATCTGTG 1464
Db 81 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100
QY 1465 ACCGCACTAATTTATCTCAAGAACACACACACATCTGAGATAAATATATCCATGACATCTTT 1524
Db 101 ThrProLeuIleIleTyrGluGluThrThrAspIleIlePheAsnIleHisAspIlePhe 120
QY 1525 CATGTTTTTCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTGGCTCTGAAATCAAA 1584
Db 121 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 140
QY 1585 ACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGAAAGCAATATATAACGA 1644
Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 160
QY 1645 TCCAGTGGTGGCTGCTCTCAAGTGTCTCAAGTGTCTCTATCAAGAGAGAGATAGCA 1704
Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 180
QY 1705 ATTACCAAGTGTGAATGGGAAGTTCTTGGCCGCAATGGATCTAATATCAAGTTGATGAA 1764
Db 181 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 200
QY 1765 GTCAGAAAGCTGTATATTTTGAAGGCAACAAGACTCCCTTTTAGAGCATCACCTGTAC 1824
Db 201 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 220
QY 1825 GTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAT 1884
Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240
QY 1885 TCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAATAGTAAACCAGAGAAT 1944
Db 241 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 260
QY 1945 CCACACTGTGTGCTTCTTACAAAGCTATCAAGCTATCAAGTCTGAAGATGACCCAACTGCAAAACA 2004
Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 280
QY 2005 AAGCAATTTGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTCTGACTATCTACTCTCCA 2064
Db 281 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300
QY 2065 GAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTCAAGCTT 2124
Db 301 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320
QY 2125 CATGATCTACAGCTCGAAAGAAATATCTTACTGTGCTGCTTTCATATATGGTGGCTCTCAG 2184
Db 321 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 340
QY 2185 GTG 2187
Db 341 Val 341

RESULT 8
US-09-976-674-13
; Sequence 13, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Q1, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
```

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 358  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-13

Alignment Scores:  
Pred. No.: 8,51e-312 Length: 358  
Score: 339.00 Matches: 339  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 32.9% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-13 (1-358)

```
QY 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATGGAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGACGGTAT 333
Db 21 GluGluAsnLeuGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGTGCGGATACAGAAATATCATGGCTACATCATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGACCCACATGATTTTCATGTTTGTGAAGAGAGAAATGATCCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGCATGCTGCTGAGACAGACAGAAATACACTGTTTATCT 513
Db 81 AspArgIleTyrIleuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAATATCCCAAACTATCAATACAGCAGCAGCTTAAATGCTCTCTTGAAGCCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGACACATCGACATGGAATGTAATTCGAGAAGAAAGAACTATTAAGA 633
Db 121 AspleuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACGCATTCGAAACAGTCGGAATTCCTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATACGTAAGAAAGATGGAGGCCACCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCCCAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTAATGCCCCCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCTAACACAGAGAAGAAAGGAGACTCACTTATGTGCACATAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnLeu 240
QY 934 GCCAATCATGAGAGATGCCAGATCAGCTGAGCTCGCTACCTTTGTTCTCCAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTGTAGATATCTGGCTATGTTGGTGTCTCAAAAGCTGAACAACTCCAGTGTGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
```

```
QY 1054 AAAATCTTAGAATCTATATGAAGAAAATGATGAATCTGAGGTGAAATATTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTATGTTGGAAACAAGAGGGCAGATTTCATCCGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGAAAGGAAGG 1230
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArg 339
RESULT 9
US-10-070-464-7
; Sequence 7, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-7
```

Alignment Scores:  
Pred. No.: 5.89e-239 Length: 360  
Score: 262.00 Matches: 262  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 25.4% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-7 (1-360)

```
QY 1756 GTTGATGAAGTCAGAAAGGCTGGTATATTTCAAGGCACAAAGACTCCCTTTAGAGCAT 1815
Db 99 ValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHis 118
QY 1816 CACCTCTAGTAGTCAGTTACGTAAATCCTGGAGAGGTGCAAGGCTGACTGACCGTGGC 1975
Db 119 HisLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGly 138
QY 1876 TACTCACATTCCTGCTGCATCAGTCAGCAGCTGCTCTTTATAGTAGTATAGTAAC 1935
Db 139 TyrSerHisSerCysCysIleSerGlnHisCysAspPheIleSerLysTyrSerAsn 158
QY 1936 CAGAGAATCCACACTGTGCTCTCTTACAGCTATCAAGTCTCTGAAGATGCCCAACT 1995
Db 159 GlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThr 178
QY 1996 TGCAAAACAAAGAAATTTTGGGCCACCATTTTGGATTCAGCAGGCTCTCTCTGACTAT 2055
Db 179 CysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyr 198
QY 2056 ACTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCGATTTCATTGTATGGGATGCTC 2115
Db 199 ThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeu 218
QY 2116 TACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTTCATATATGTT 2175
```

Db 219 TyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGly 238  
QY 2176 GGTCTCAGGTCAGTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAT 2235  
Db 239 GlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsn 258  
QY 2236 ACCCTAGCTCTTAGGTTATGTTGGTTAGTATAGACAACAGGGGATCTCTGCACGA 2295  
Db 259 ThrLeuAlaSerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArg 278  
QY 2296 GGGCTTAAATTTGAAGCGCTTTAAATATAAATGAGTCAAAATGAGCATGACAG 2355  
Db 279 GlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluLeuAspGln 298  
QY 2356 GTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGC 2415  
Db 299 ValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspValGly 318  
QY 2416 ATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTCAAT 2475  
Db 319 IleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAsp 338  
QY 2476 ATCTTCAGGGTGTATTGCTGGGGCCAGTCCACTCTGTGATCTTCTATGATACACGA 2535  
Db 339 IlePheArgValAlaIleAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGly 358  
QY 2536 TACACG 2541  
Db 359 TyrThr 360  
RESULT 10  
US-09-976-674-9  
; Sequence 9, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976, 674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-9  
Alignment Scores:  
Pred. No.: 3.18e-216 Length: 241  
Score: 238.00 Matches: 238  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 23.1% Indels: 0  
DB: 2 Gaps: 0  
US-10-825-632-2 (1-3120) x US-09-976-674-9 (1-241)  
QY 214 ATGCAGCAGCAATGGAACAGACAGCTGGGTGTTGAGATATTTGAAACTGCGACTGT 273  
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGAGATATTGAAATCACAGGATCGGCCTTAAATTTGGAGCCCTTTTATGTTGACGGTAT 333  
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCCTGGAGTACAGTTAAAGAGCTGCTGGCGATACACAGAAATATCATCGGTACATGATG 393

Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 GCTAAGGCACCACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453  
Db 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 GACAGATCTATTACCTTGGCATGCTGCTGAGAACAGAGAAATACACTGTTTATTTCT 513  
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
QY 514 GAAATTTCCAAAACATCAATAGACGAGCAGTCTTAATGCTCTCTCTTTGGAAGCCTCTTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 CATCTTTTTCAGCAACACATGGACTAGGAATGATTTCTCGAGAGAGAACTATTAAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
QY 634 GAAAGAAAACGATTCGACAGTCGGAATGCTTCTTACGATTATCACCAAGGAAGTGA 693  
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCACTAAAGATCGAGGGCCACAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCAAAATATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873  
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCACAAAT 927  
Db 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsn 238  
RESULT 11  
US-09-976-674-17  
; Sequence 17, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976, 674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 194  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-17  
Alignment Scores:  
Pred. No.: 3.52e-164 Length: 194  
Score: 183.00 Matches: 183  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 17.8% Indels: 0  
DB: 2 Gaps: 0  
US-10-825-632-2 (1-3120) x US-09-976-674-17 (1-194)  
QY 214 ATGCAGCAGCAATGGAACAGACAGCTGGGTGTTGAGATATTTGAAACTGCGACTGT 273

Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGAGAAATATTGAATCACAGGATCGGCCTAAATTCGAGCCCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTy-ValGluArgTyr 40  
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACAGAAAAATATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTy-HisGlyTy-MetMet 60  
QY 394 GCTAAGGACACCATGATTTTCATGTTTGTGAGAGAAATGATCCAGATGGACCTCATTC 453  
Db 61 AlaLysAlaProHisAspPheMetPheValLysAspAsnAspProAspGlyProHisSer 80  
QY 454 GACAGATCTATTACCTTGCATGCTGCTGAGACAGAGAAATACACTGTTTATCT 513  
Db 81 AspArgIleTyTyLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTySer 100  
QY 514 GAATTTCCCAAACTATCAATAGACGACGAGCTTAAATGCTCTCTTGGAAAGCTCTTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTCAGCAACACTGGAATGATGGAATGATTTCTCGAGAGAAAGAACTATTAA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
QY 634 GAAGAAGAACGCAATGGAACAGTCGGAATGCTTTCACGATTATCACCAAGGAAGTGA 693  
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTTCAAGCGGTAGTGAATTTTATCATGTAAGATGAGAGGCCACAAAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTACGCCAA 762  
Db 181 PheThrGln 183

RESULT 12  
US-09-794-236-4  
; Sequence 4, Application US/09794236  
; Patent No, 6337069  
; GENERAL INFORMATION:  
; APPLICANT: Grouzmann, Eric  
; APPLICANT: Lacroix, Jean-Silvain  
; APPLICANT: Monod, Michel  
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis  
; FILE REFERENCE: 81985/276823  
; CURRENT APPLICATION NUMBER: US/09/794,236  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 310  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-794-236-4

Alignment Scores:  
Pred. No.: 1,37e-157 Length: 310  
Score: 176.00 Matches: 176  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 17.1% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-794-236-4 (1-310)

QY 2332 GGTCAAATGAATTAACATCAGGTGAGAGCACTCCAAATATCTAGCTTCTCCATATGAT 2391  
Db 135 GlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyLeuAlaSerArgTyrAsp 154  
QY 2392 TTCATTGACTTAGATCGTGTGGGCATCCACGCTGGTCCTATGAGGAGTACCTCTCCCTG 2451

Db 155 PheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyTyLeuSerLeu 174  
QY 2452 ATGCATTATATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACT 2511  
Db 175 MetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThr 194  
QY 2512 CTGTGGATCTCTATGATACAGGATACAGGACGTTATATGGTCACCTGACCCAGAT 2571  
Db 195 LeuTrpIlePheTyAspThrGlyTyThrGluArgTyRwecGlyHisProAspGlnAsn 214  
QY 2572 GAAACAGGCTATTACTTAGGATCTGTGGCATCGCAAGCAGAAAAAGTTCCCTCTGGAACA 2631  
Db 215 GluGlnGlyTyTyLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluPro 234  
QY 2632 AATCGTTACTGCTCTTACATGTTCTCTGGATGAGAATGTCCATTTTGCACATACCAAGT 2691  
Db 235 AsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSer 254  
QY 2692 ATATTACTAGTGTCTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCCTCAG 2751  
Db 255 IleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyProGln 274  
QY 2752 GAGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACATTATGAACATCTCTTTTGCAC 2811  
Db 275 GluArgHisSerIleArgValProGluSerGlyGluHisTyGluLeuHisLeuLeuHis 294  
QY 2812 TACCTTCAAGAAACCTTGGATCCAGTATTGCTGCTCTTAAAGTGATA 2859  
Db 295 TyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 310

RESULT 13  
US-10-070-464-3  
; Sequence 3, Application US/10070464  
; Patent No, 6881564  
; GENERAL INFORMATION:  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 310  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-070-464-3

Alignment Scores:  
Pred. No.: 1,37e-157 Length: 310  
Score: 176.00 Matches: 176  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 17.1% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-3 (1-310)

QY 2332 GGTCAAATGAATTAACATCAGGTGAGAGCACTCCAAATATCTAGCTTCTCCATATGAT 2391  
Db 135 GlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyLeuAlaSerArgTyrAsp 154  
QY 2392 TTCATTGACTTAGATCGTGTGGGCATCCACGCTGGTCCTATGAGGAGTACCTCTCCCTG 2451  
Db 155 PheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyTyLeuSerLeu 174

QY 2452 ATGGCATTAAATGAGAGGTTCAGATATCTTCAGGGTTCCTATTCTCGGGCCCCAGTCACT 2511  
Db 175 MetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThr 194  
QY 2512 CTGTGATCTTCTATGATACAGGATACACGGAAGTTATATGGTTCACCTGACACAGAT 2571  
Db 195 LeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsn 214  
QY 2572 GACAGGGCTATTACTAGATCTGTGCCATGCTCAAGCAGAGAAAGTTCCCTCTGAACCA 2631  
Db 215 GluGlnGlyTyrTrpLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluPro 234  
QY 2632 AATCGTTTACTGCTTACATCGTTTCTCGATGAGAAATGTCATTTTGCACATACAGT 2691  
Db 235 AsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSer 254  
QY 2692 ATATTACTAGTTTTTTAGTGGCTGGAAGCCATGATTTACAGATCTATCTCAG 2751  
Db 255 IleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGln 274  
QY 2752 GAGACACAGCATAGAGTTCCTGAATCGGGAGAACATTTATGAATCTCTTTTGCAC 2811  
Db 275 GluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHis 294  
QY 2812 TACCTTCAAGAAACCTTCGATCAGTATTGTCTCTTAAAGTGATA 2859  
Db 295 TyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 310  
RESULT 14  
US-09-976-674-15  
; Sequence 15, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-15  
Alignment Scores:  
Pred. No.: 2,23e-72 Length: 108  
Score: 86.00 Matches: 86  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 8.3% Indels: 0  
Gaps: 2  
Db:  
US-10-825-632-2 (1-3120) x US-09-976-674-15 (1-108)  
QY 214 ATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTGAACTCGGACTGT 273  
Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGAGATATTGAATTCAGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCCTGGAGTCAGCTTAAAGCTCTTGGCGATACAGAAATATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60


QY 394 GCTAAGGCACCACATGATTTTCATGTTTGAAGAGGAATGATCCAGATGGACCTCATCA 453  
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 GACAGAATCTATTACCTT 471  
Db 81 AspArgIleTyrTyrLeu 86  
RESULT 15  
US-09-621-976-5647  
; Sequence 5647, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5647  
; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: SIGNAL  
; NAME/KEY: -39...-1  
; LOCATION: -39...-1  
US-09-621-976-5647  
Alignment Scores:  
Pred. No.: 9,24e-30 Length: 54  
Score: 41.00 Matches: 41  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 4.0% Indels: 0  
Gaps: 2  
Db:  
US-10-825-632-2 (1-3120) x US-09-621-976-5647 (1-54)  
QY 2602 ATGCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTACTGCTTACATGGTTTCTG 2661  
Db 1 MetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeu 20  
QY 2662 GATGAGAATGTCATTTTGCACATACATATATTACTGAGTATTTTATGTCAGGCTGGA 2721  
Db 21 AspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGly 40  
QY 2722 AAG 2724  
Db 41 Lys 41  
Search completed: May 2, 2006, 03:55:26  
Job time : 128.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: May 2, 2006, 04:41:36 ; Search time 85.9 Seconds  
(without alignments)  
3035.221 Million cell updates/sec

Title: US-10-825-632-2  
Perfect score: 1030  
Sequence: 1 aagtgctaagctccgagg.....agaattactataaaaaaaa 3120

Scoring table:   
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 60.0, Fgapext 7.0  
Delop 60.0, Delext 60.0

Searched: 1867569 seqs, 417829326 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3572506

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:  
-MODE=frame+n2p.model -DEV=xlp  
-Q/abss/ABSSWEB\_spool/US10825632/runat\_01052006\_110247\_4670/app\_query.fasta\_1  
-DB=Published Applications AA Main -QFMT=fastan -SUFFIX=oligo.n2p.rapbm  
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abss02p  
-USER=US10825632 @CGN 1 1 307 @runat\_01052006\_110247\_4670 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database : Published Applications AA Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	882	85.6	882	3	US-09-976-674-1	Sequence 1, Appli
2	882	85.6	882	4	US-10-054-776-2	Sequence 2, Appli
3	882	85.6	882	4	US-10-170-789-38	Sequence 38, Appli
4	882	85.6	882	4	US-10-311-035-9	Sequence 9, Appli
5	882	85.6	882	4	US-10-072-012-622	Sequence 622, App
6	882	85.6	882	4	US-10-415-122-6	Sequence 6, Appli
7	882	85.6	882	4	US-10-825-632-1	Sequence 1, Appli
8	882	85.6	882	5	US-10-982-512-1	Sequence 1, Appli
9	655	63.6	658	3	US-09-976-674-19	Sequence 19, Appli
10	655	63.6	658	5	US-10-982-512-19	Sequence 19, Appli
11	655	63.6	661	3	US-09-976-674-11	Sequence 11, Appli

12	655	63.6	661	5	US-10-982-512-11	Sequence 11, Appli
13	655	63.6	690	3	US-09-976-674-7	Sequence 7, Appli
14	655	63.6	690	5	US-10-982-512-7	Sequence 7, Appli
15	607	58.9	613	3	US-09-976-674-21	Sequence 21, Appli
16	607	58.9	613	5	US-10-982-512-21	Sequence 21, Appli
17	341	33.1	465	4	US-10-825-632-5	Sequence 5, Appli
18	339	32.9	358	3	US-09-976-674-13	Sequence 13, Appli
19	339	32.9	358	5	US-10-982-512-13	Sequence 13, Appli
20	262	25.4	360	4	US-10-825-632-7	Sequence 7, Appli
21	238	23.1	241	3	US-09-976-674-9	Sequence 9, Appli
22	238	23.1	241	5	US-10-982-512-9	Sequence 9, Appli
23	194	18.8	580	4	US-10-275-505-2	Sequence 2, Appli
24	194	18.8	580	6	US-11-140-224-2	Sequence 2, Appli
25	186	18.1	883	4	US-10-072-012-621	Sequence 621, App
26	183	17.8	194	3	US-09-976-674-17	Sequence 17, Appli
27	183	17.8	194	5	US-10-982-512-17	Sequence 17, Appli
28	176	17.1	310	3	US-09-993-959-4	Sequence 4, Appli
29	176	17.1	310	4	US-10-825-632-3	Sequence 3, Appli
30	151	14.7	193	4	US-10-264-237-2115	Sequence 2115, Ap
31	86	8.3	108	3	US-09-976-674-15	Sequence 15, Appli
32	86	8.3	108	5	US-10-982-512-15	Sequence 15, Appli
33	79	7.7	129	3	US-09-764-891-3564	Sequence 3564, Ap
34	19	1.8	19	3	US-09-976-674-59	Sequence 59, Appli
35	19	1.8	19	5	US-10-982-512-59	Sequence 59, Appli
36	17	1.7	18	4	US-10-825-632-22	Sequence 22, Appli
37	15	1.5	16	4	US-10-825-632-23	Sequence 23, Appli
38	13	1.3	432	4	US-10-237-271-4	Sequence 4, Appli
39	13	1.3	518	3	US-09-976-674-25	Sequence 25, Appli
40	13	1.3	518	4	US-10-729-807-16	Sequence 16, Appli
41	13	1.3	518	5	US-10-982-512-25	Sequence 25, Appli
42	13	1.3	819	3	US-09-976-674-37	Sequence 37, Appli
43	13	1.3	819	3	US-09-976-674-39	Sequence 39, Appli
44	13	1.3	819	5	US-10-982-512-37	Sequence 37, Appli
45	13	1.3	819	5	US-10-982-512-39	Sequence 39, Appli

ALIGNMENTS

RESULT 1  
US-09-976-674-1  
; Sequence 1, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Oi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976, 674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-1

Alignment Scores:  
Pred. No.: 0  
Score: 882.00  
Percent Similarity: 100.0%  
Best Local Similarity: 100.0%  
Query Match: 85.6%  
DB: 3  
Length: 882  
Matches: 882  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-1 (1-882)

QY 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAACCTGCGACTGT 273  
|||||



Db 1 MetAlaAlaMetGluThrGluGluLeuGlyValGluIlePheGluThrAlaAspCys 20  
Qy 274 GAGGAGATATGTAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
Qy 334 TCTGGAGTCAGCTTAAAGCTCTGTTGCCGATACACAGAAAATATATCATGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuIleLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
Qy 394 GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453  
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
Qy 454 GACAGAATCTATTACCTTGGCATGCTCTGTGTGAGAACAGAGAAAAATACACTGTTTATTCT 513  
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
Qy 514 GAAATTTCCAAAACATATCAATAGACGACGAGTCTTAATGCTCTCTTGAAGCCTCTTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
Qy 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATATCTCGAGAGAGAAAGAACTATTAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
Qy 634 GAAAGAAACCATTTGAAACAGTCGGAATTTCTTACGATTATCCACAGGAGAGTGA 693  
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
Qy 694 ACNTTTCNTGTTCAAGCCGCTAGTGGNATTTATCAGTAAAGATGGAGGCCACAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyTyrHisValLysAspGlyGlyProGlnGly 180  
Qy 754 TTTACGCCAACCTTTTAAGCCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
Qy 814 GATCCAAATTTATCCCGCTGATCCAGACTGGATGCTTTTATATAGTACAAACGATATT 873  
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
Qy 874 TGGATATCTAAACATCGTAACACAGAGAAAGAGACTCATTATGTCCCAATAGACTA 933  
Db 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240  
Qy 934 GCCAACATGGGAAGATGCCAGATCAGCTGGAGTCGCTTACCTTTTGTCTCCAAAGAGAA 993  
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260  
Qy 994 TTTGATAGATATCTGGCTATTGCTGTGTCCAAAAGCTGAAAACACTCCCAGTGGTGGT 1053  
Db 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280  
Qy 1054 AAAATTTCTAGAATTTCTATATGAAGAAATGATGATCTGAGTGGAAATTTATTCATGTT 1113  
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300  
Qy 1114 ACNTCCCTATGTTGGAACACAGGGCAGATTCATTTCCGTTATCTCTAAACAGGTACA 1173  
Db 301 ThrSerPrometLeuGluThrArgArgAlaAspSerPheArgTyrProLysTrpGlyThr 320  
Qy 1174 GCAATCTCAAGTCATTTTAAAGTCGACGAAATTAATGATTTGATGCTGAAGGAGATC 1233  
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
Qy 1234 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGTAGATTTCTATTGAAAGGAGTTGAA 1293  
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
Qy 1294 TATATTGCCAGAGTCGATCGACTCTCTGAGGAGAAATATGCTTGTGCTCATCTACTAGAT 1353  
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380

Qy 1354 CGCTCCCGACTCGCCTACAGATAGTGTGATCTCACTGAATTTATTTATCCAGTAGAA 1413  
Db 381 ArgSerGlnThrArgLeuGluIleValLeuIleSerProGluLeuPheIleProValGlu 400  
Qy 1414 GATGATGTTATGGAAAGGACAGACTCATTTGAGTCAAGTCCCTGATCTGTGTGACGCCACTA 1473  
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
Qy 1474 ATTATCTATGAAGAAACACACAGACATCTCGATAAATATCCATGACATCTTTCATGTTTTT 1533  
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
Qy 1534 CCCCAAGATCACGAAGAGGAAATTTGAGTTTATTTTTCCTCTGAATTCGAAACAGGTTTC 1593  
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
Qy 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT 1653  
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
Qy 1654 GGGCTGCTGCTCCCAAGTGTCTCAAGTGTCTCTATCAAAGAGAGATAGCAATTTACCAGT 1713  
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500  
Qy 1714 GGTCAATGGGAAGTCTTGGCCGCGCATGGATCTAATATCCAGTTTGATGAAAGTCAGAGG 1773  
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520  
Qy 1774 CTGCTATATTTTGAAGCCACCAAGACTCCCTTTTAGAGCATCCTCTAGTGTAGTCACT 1833  
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540  
Qy 1834 TACGTAAATCTCGAGAGGTGACAGCTGACTGACCGTGGCTACTCATCTTCTGCTGC 1893  
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560  
Qy 1894 ATCAGTCAGACCTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGT 1953  
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
Qy 1954 GTGTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCAACTGTGCAAAACAAAGAAATTT 2013  
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
Qy 2014 TGGGCCCACTTTTGGATTCCAGCAGGTCTCTTCTCTGACTATCTCTCCAGAAATTTTC 2073  
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620  
Qy 2074 TCTTTTGAAAGTACTACTGGATTACATTTGATGGATGCTCTCAAGACCTCATGATCTA 2133  
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
Qy 2134 CAGCTCGAAAGAAATATCTACTGTGTTTTCATATATGTTGGTGGTCTCCTCAGTGCAGTTG 2193  
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660  
Qy 2194 GTGAATATCGTTTAAAGGAGTCAAGTATTTCCCTTGAATACCTAGCTCTCTAGGT 2253  
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680  
Qy 2254 TATGTGTTCTAGTATAGACAACAGGGCATCTCTCACCAGGGCTTAAATTTGAGGC 2313  
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700  
Qy 2314 GCCTTTAAATATATAAATGGGTCAATAGAAATTTGACATCAGTCTGGAAGGACTCAATAT 2373  
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720  
Qy 2374 CTAGCTCTCGATATGATTTTCATTTAGATTCGTGTGGGCATCCACGGCTGCTCTAT 2433  
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740

2434 GAGGATACCTCCCTGATGCATTAAATGACAGAGGTGAGATATCTTCAGGTTGCTATT 2493  
 Db |||||TyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760  
 2494 GGTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTATTATG 2553  
 Db |||||AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluA-GTyrMet 780  
 2554 GGTACCCCTGACAGAAATGAACAGGGCTATTACTTATGATACAGGATACAGGATACAGGAA 2613  
 Db |||||GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800  
 2614 AGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGATGTC 2673  
 Db |||||LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820  
 2674 CATTTTGCATACACAGTATATTACTGAGTTTTTGTAGTGAGGGCTGGAAAGCCATATGAT 2733  
 Db |||||HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840  
 2734 TTACAGATCTATCTCAGGAGACACAGACATAAGAGTTCTCGAATCGGAGAACATATTAT 2793  
 Db |||||LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860  
 2794 GAATGATCTTTTGCACACTACCTTCAAGAAACCTTTGGATCACGTTATGTGCTCTAAAA 2853  
 Db |||||GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880  
 2854 GTGATA 2859  
 Db |||||ValIle 882

RESULT 2  
 US-10-054-776-2  
 ; Sequence 2, Application US/10054776  
 ; Publication No. US20030165818A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mark Robert Edbrooke  
 ; APPLICANT: Alan Peter Lewis  
 ; TITLE OF INVENTION: NOVEL PROTEIN  
 ; FILE REFERENCE: QG1042US  
 ; CURRENT APPLICATION NUMBER: US/10/054,776  
 ; CURRENT FILING DATE: 2002-01-23  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 882  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-054-776-2

Alignment Scores:  
 Pred. No.: 0 Length: 882  
 Score: 882.00 Matches: 882  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 85.6% Indels: 0  
 DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-054-776-2 (1-882)

214 ATGGCAGCAGCAATGGAACAGAACAGAGCTGGTGGTTGAGATATTTGAAACTGCGAGTGT 273  
 Db |||||MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
 274 GAGGAGAAATATGAATCACAGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333  
 Db |||||GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
 334 TCCTGGATGAGCTTAAAGAGCTCTTGGCGATACAGAAATATCATGGCTACATGATG 393  
 Db |||||SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

394 GCTAAGGCCACCATCATTTTCATGTTTCTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453  
 Db |||||AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
 454 GACAGAATCTATTACCTTCCCATGTCGTGAGAACAGAGAAAATACACTGTTTATTCTTCT 513  
 Db |||||AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
 514 GAAATTTCCCAAACTATCAATAGACAGCAGCTTAATGCTCTCTTCCGAAAGCCCTCTTTTG 573  
 Db |||||GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
 574 GATCTTTTTCAGGCAACACTGACTATGGAATGATTCTCGAGAAGAAGAACTATTAAAGA 633  
 Db |||||AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
 634 GAAAGAAAACGCAATTGGAAACAGTCGGAATTCCTTACGATTATTCACAAAGGAAGTGGGA 693  
 Db |||||GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
 694 ACATTTCTGTTTCAAGCCGTTAGTGGATTTTATTCACGTAAAAAGATGGAGGCCCAAGGA 753  
 Db |||||ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGATG 813  
 Db |||||PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
 814 GATCCAAAATTATGCCCGCTGATCCAGACTCGATTGCTTTTATACATAGCAACGATATT 873  
 Db |||||AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
 874 TGGATATCTAACATCGTAACACAGAGAAAGAGAGACTCATTATGTGCACAATGAGCTA 933  
 Db |||||TrpIleSerAsnIleValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeu 240  
 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAA 993  
 Db |||||AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260  
 994 TTTGATAGATATTCTGGCTATTGGTGTGCCAAGCTGAAAGCTGAAACACTCCAGTGGTGGT 1053  
 Db |||||PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
 1054 AABATTTCTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATATTATCATGTT 1113  
 Db |||||LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300  
 1114 ACATCCCTCTATGTTGGAACACAGAGGGGAGATTTCATTCGGTTATCTCTAAACAGGTACA 1173  
 Db |||||ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
 1174 GCAAACTCTAAAGTCATCTTTTAAGATGTGAGAAAATAAGATTGATGCTGAGGAAGGATC 1233  
 Db |||||AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
 1234 ATAGATGTCATAGTAAGGAACCTAATTCACCTTTTGAGATTCTATTGAGGAGGTTCGAA 1293  
 Db |||||IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
 1294 TATATTTGCCAGAGCTGATCGACTCTCTGAGGGAATAATATGCTTGGTCCATCTCTAGAT 1353  
 Db |||||TyrIleAlaArgAlaGlyTyrTrpProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380  
 1354 CGCTCCAGAGCTCCGCTTACAGATAGTGTGATCTCACTCGAATATTATTATCCAGTAGAA 1413  
 Db |||||ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
 1414 GATGATGTATTGGAAGGACAGACACTCATGTAGTCAGTCCTGATTCGTGTGAGCCACTA 1473  
 Db |||||AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
 1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533

```
Db 421 ileiletyrGluGluThrThrAspIletrPileAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCAAAGTCACGAAGAGGAAATGAGTTATTTTTCCTCTGAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluGluPhePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 GTCATTATATCAAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysIleArgSerSerGly 480
QY 1654 GGGCTGCTCTCCCAAGTGATTTCAAGTGCTCTATCAAAAGAGAGATAGCAATACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGAGAGTTCTTGCCCGGCATGGATCTATATCAAGTTGATGAAGTCAGAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTACCAAGCTATCAAGTCTCTGAAGATGACCAACTTCGCAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGATTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTCGATTTACATGTATGGATGGCTCTACAGCTCTATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTTGGGAAGAAATATCTACTGCTGCTGTTCATATATGGTGTCTCAGGTGCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTCAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCPAGCCTCTCTAGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGGTTGTAGTAGACACACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATGGGTCAAAATAGAAATTCAGCATCAGGTGGAGGACTCCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTCAATTGACTTAGATTCGTGGGCATCCACGGCTGGTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyLysTrpSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCATTATGACAGAGGTACAGATCTTCAGGGTTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACCTGTGGATCTTCATGATACAGGATACACGGAAGGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCCCTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCTGCAACAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCACATACCAGTATATATCTAGTGTCTTTTGTAGTGAGGCTGGAAAGCCATATCAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCTCAGGAGAGACACAGATCAAGAGTTCTCTGAATCGGAGAAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACTGATCTTTTGCATACCTCTCAAGAAACCTTGGATCACGTATTGCTGCTCTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
RESULT 3
US-10-170-789-38
; Sequence 38, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Ribermann, Rosana
; APPLICANT: Cuttiti6 , Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Melich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
```

```
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-789-38

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-170-789-38 (1-882)

QY 214 ATGCGACGACAAATGAAACAGAACAGCTGGTGTGAGATATTGAAACATGCGGACTGT 273
DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATCAGAGGATCGGCTAAATTTGGAGCCCTTTTATTTGTTGACCGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGGCTCATCATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACACATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA 453
DB 61 AlalysAlaProHisaspPheMetPheValLysArgAsnAppProaspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGCATGCTCGGTGAGACAGAGAAATACACTGTTTATTCT 513
DB 81 AspArgIleTyrIleuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAAATCATCAATAGACGACGAGCTTTAATGCTCTCTTGGAAACCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGGCAACACTGGACTATGGAATGTTATCTCGAGAGAGAGAACTATTAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAACCCATGGAAACAGTCGGAAATGCTTCTTACGATTATCCAAAGGAGGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGATGGAGGCGCCCAAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAACTAGTTGTGCCAACATACGGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATTGCCCGCTGATCCAGACTGGGATTGCTTTTATACATAGCAACGATATT 873
DB 201 AspProLysLeuCysProAlaAspProaspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAACCCAGAGAAAGAGACTCACTTATGTGCAATAGACTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTTCTCCAGAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGTCCAAAGCTGAAACAACTCCACAGTGGTGGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTCTTAGAATCTATATGAAGAAATATGATGAATCTGAGGTGGAAATATTTCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTATGTTGAAACAAAGGAGGCGAGATTCATTCCTTATTCCTAAACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGATGCTGAAGGAGGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCNTAGATAAGAACTAATTAACCTTTTGAGATCTTATTTGAGGAGTTGAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGACTCGATGGACTCTGAGGGAATATGCTTGGCTCATCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAGGAGAGACTCATGTAGTCAGTGCCTGATCTGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCACGAGAGGAAATGAGTTTATTTTGGCTCTGATGATCAAAACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATCAAAATTTACATCTATTTTAAAGAAAGCAATATAAATACGATCCAGTGGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAAGT 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTTGCGCGGATGATCTAATATCCAAAGTGTGATGAAGTCAGAGG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGCCACCAAGACTCCCTTTTAGACATCACCTGACTAGTCTAGT 1833
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACCTTAATCTCGAGAGGTCACAGCGTACTGACCGTGGCTACTCACATCTTCTGCTGC 1893
DB 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
```

```
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATTAAGTAAGTATAGTAAACCAAGAATCCACACTGT 1953
Db |||||
QY 561 IISerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db |||||
QY 1954 GTGTCCTTTTACAAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCABAACAAAGGAATTT 2013
Db |||||
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCAATTTTGGATTTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTTC 2073
Db |||||
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTCGATTACATTGTATGGGATGCTCTCAAGACCTTCATGATCTA 2133
Db |||||
QY 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCTGGAAGAAATATCTACTGTCTGTTCATATATATGGTGGTCTCAGGTGCAGTTG 2193
Db |||||
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT 2253
Db |||||
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTGTAGACACACAGGGATCTCTGTCCAGAGGCTTAAATTTGAAGGC 2313
Db |||||
QY 681 TyrValValValLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLysPheGluGly 700
QY 2314 GCCTTTAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATAT 2373
Db |||||
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLysGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTCAATTCATTCATGATCGTGTGGCATCCAGCTGGTCTAT 2433
Db |||||
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATGACAGGTCAGATATCTTCAGGTTGCTATT 2493
Db |||||
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCACTGCTCTGTGGATCTCTTATGATACAGATACAGGAACTGTATATG 2553
Db |||||
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTGACACAGAAATGAACAGGCTTATCTTAGATCTGTGGCCATGCAAGCAGAA 2613
Db |||||
QY 781 GlyHisProAspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGTGGATGAGAATGTC 2673
Db |||||
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCATACACAGTATATTACTAGTGTTTTGTAGTGGGCTGGAAAGCCATATGAT 2733
Db |||||
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCTCAGGAGACACACAGCATAAGAGTTCTGAATCGGAGAACATTAAT 2793
Db |||||
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACTCACTCTTTTGCACCTACCTTCAAGAAACCTTCGATCAGCTATCTGCTCTCAAA 2853
Db |||||
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db |||||
QY 881 ValIle 882
```

RESULT 4

US-10-311-035-9

; Sequence 9, Application US/10311035

; Publication No. US20040023243A1

```
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depodriya
; APPLICANT: KERNEY, Adam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Processes
; FILE REFERENCE: FI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1
US-10-311-035-9
```

Alignment Scores:

Pred. No.:	0	Length:	882
Score:	882.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	85.6%	Indels:	0
DB:	4	Gaps:	0

US-10-825-632-2 (1-3120) x US-10-311-035-9 (1-882)

QY	214	ATGGCAGCAGCAATCGAAACAGACACGCTGGCTTCAGATATTTGAACTCGGACTGT	273
Db	1	MetAlaAlaAlaMetGlnThrGluGluLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGAGATATTGAATCAAGATCGCGCTTAAATTGAGCGCTTTTATGTTGACGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCTTGAGTCACTTAAAAAGCTGCTGCGCATACCAAGAAAATATCATGGCTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAATCTATTACCTTGCCTATCTCTGGTGAACACAGAGAAAATACACTGTTTATCT	513

Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
QY 514 GAAATCCCAAACTATCAATAGACGACGAGTCTTAATGCTCTCTGGAAGCCTCTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaValIleuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTTCAGGCAACACTCGACTATGGAATGATTTCTCGAGAAGAACAATTAAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140  
QY 634 GAAAGAAACCATTTGAACAGTCGGAATTCCTTCTACGATTATCAACCAAGGAAGTGA 693  
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGAGGGGCCACAAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTAGCCCAACACCTTTAAAGCCCAATCTAGTGGAACTAGTGTGCCCAACATAGGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCCAAATTAATGCCCCGCTGATCCAGACTGGATTGCTTTTATATACATAGCAACGATAT 873  
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTAAATCGTAACACAGAGAAGAGAGACTCACTTATGTGCAACATGAGCTA 933  
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240  
QY 934 GCCAACATGGAAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAAGAGAA 993  
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260  
QY 994 TTTGATAGATATTCGGCTATTTGGTGTGCCAAAAGCTGAAACAACTCCCAAGTGGTGG 1053  
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATCTTAGAATCTATAGAANAATGATGAATCTGAGGTGGGAATTTATTCATGTT 1113  
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300  
QY 1114 ACATCCCTATGTTGGAACAAGAGGGCAGATTCATTCGTTATCTCTAAACAGGTACA 1173  
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATGATGCTGGAAGAGGATC 1233  
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleuMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGAGATCTTATTTGAAGAGTGA 1293  
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
QY 1294 TATATGCGCAGCTGGATGACCTTCGAGGGAATAATGCTTGGTCCATCTCTACTAGAT 1353  
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380  
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTTCATCTCACTCGAATTTATTTATCCAGTAGAA 1413  
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATGATGTTATGGAAGCGCAGACTCATTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473  
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533  
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
QY 1534 CCCCAGTCAAGAGGAATTTAGTTTATTTTTCCTCTGAAATGCAAAACAGGTTTC 1593  
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460

QY 1594 CGTCATTTTATACAAAATTACATCTATTATTAAGGAAAGCAAAATATAAAGCATCCAGTGT 1653  
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
QY 1654 GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTCAAAAGAGAGAGATAGCAATTAACAGT 1713  
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500  
QY 1714 GGTGAATGGGAAGTCTTGGCCGCGATGATCTAATATCAAGTTGATGAAGTCAGAGG 1773  
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520  
QY 1774 CTGGTATATTTGAAGGACCAAGACTCCCTTTAGACATCACCCTGACCTAGTCACT 1833  
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540  
QY 1834 TACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893  
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560  
QY 1894 ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953  
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
QY 1954 GTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAAACAAAGAAATTT 2013  
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
QY 2014 TGGGCCACATTTTGGATTCAGCAGGTCTCTCTTCCTGACTATATCTCTCCAGAAATTTTC 2073  
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620  
QY 2074 TCTTTTGAAGTACTACTGCAATTTACATTTATGATGGATGCTCTACAGCTCATCATCTA 2133  
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
QY 2134 CAGCTCGAAAAGAAATCTCTACTGTCTTCATATATGTTGGTCTCAGGTGCAGTTG 2193  
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu 660  
QY 2194 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCCTCTAGGT 2253  
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680  
QY 2254 TATGTGTTGTGATAGACAACAGGGGATCTCTGTCACCGGGCTTAAATTTGAAGGC 2313  
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700  
QY 2314 GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT 2373  
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720  
QY 2374 CTAGCTTCTCATATGATTTTCATTTGATCTAGATCGTGGGCATCCACGGCTGCTCTAT 2433  
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740  
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATTCAGAGTCAAGATATCTTACGGGTGCTGCTATT 2493  
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760  
QY 2494 GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2553  
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780  
QY 2554 GGTCAACCTGACAGAAATGAACAGGGCTATTACTAGGATCTCTGGCCATGCAACAGAA 2613  
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800  
QY 2614 AAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGCTGTTTCTTGATGAGATGTC 2673  
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820



```
QY 2674 CATTTGCGACATACAGTATATTAAGTCTTTTCTAGTCCGAGCTGGAAAGCCATATCAT 2733
Db 821 HSPHeAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyProTyrAsp 840

QY 2734 TTACAGATCTATCTCTAGGAGACACAGCATTAAGTTCCTGAATCGGAGAACATAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860

QY 2794 GAATGCACTCTTTTGCACCTACCTTCAAGAAACCTTTGGATCACGATTGCTGCTCTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaAlaLeuLys 880

QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 5
US-10-072-012-622
; Sequence 622, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spvtek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Murali dhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 622
; LENGTH: 882
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-622

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-072-012-622 (1-882)

QY 214 ATGGCAGCAGCAATGGAACACAGAACAGCTGGGTGTGGAGATATTTGAAATCGCGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

QY 274 GAGGAGAATATTGAATCACAGGATCGCCTAAATTTGGAGCCTTTTATTTGTTGAGCGGTAT 333
Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40

QY 334 TCCTGGAGTCAGCTTAAAGCTGCTCCGATACCCAGAAATATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

QY 394 GCTAAGGCACACATGATTTTCATGTTGTGAGAGAGATGATCCAGATGGACCTCATCA 453
Db 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80

QY 454 CACAGAATCTATTACCTTGCCTGTCATGCTGTGTGAGAACACAGAAATATACACTGTTTATCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer 100

QY 514 GAAATTCCTCAAACTATCAATAGAGCAGCAGCTTAAATGCTCTCTTGGAAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120

QY 574 GATCTTTTTCAGGCAACACTGGAATATGGAATGTATTCGAGAGAAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140

QY 634 GAAAGAAAACGCTTTGGAACAGTCGGAATTTGCTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTCAAGCCGCTAGTGAATTTATCACGTAAAGATGGAGGGCCCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCCCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAenLeuValGluThrSerCysProAenIleArgMet 200

QY 814 GATCCAAATATGCTCCCGCTGATCCAGACTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220

QY 874 TGGATATCTAACTCGTACCAGAGAGAAAGAGACTCATTATGTCACAAATAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLysLeuThrTyrValHisAsnGluLeu 240

QY 934 CCCAAATGGAAGAAAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaIleThrPheValLeuGlnGluGlu 260

QY 994 TTTGATAGATATTCTCGCTATTGTTGGTGTCTCCAAAGCTGAAACAACTCCAGCTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280

QY 1054 AAAATCTTGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAenAspGluSerGluValGluIleIleHisVal 300
```



```
QY 1114 ACATCCCTATGTTGGAACAAGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTTACA 1173
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCTATTGGAAGGATTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATGCCAGACTGGATGGACTCTGAGGAAAAATATGCTTGGTCCATCCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGTCCCGAGACTCGCTACAGATAGTGTGATCTCACCTGAATTATTTATCCCAAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATCATGTTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerProGluLeuPheIleProValGlu 420
QY 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
Db 421 IleIleTyrGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCACGAAGGAAATTTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAATATACATCTATTTAAAGGAAGCAATATAACGATCCAGTGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAGTAGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTCAATGGGAAGTCTTTGGCCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCAATCACTCTAGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACGTAAATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGACTGCTGACTCTTTTATAGTAAGTATAGTAAACCAAGAATCCACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTACAGCTATCAAGTCTGAGATGACCCCAACTTGGCAAAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGCCACCATTTTGGATTTCAGCAGGTCCTCTCTCTGACTACTCTCCCAAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGAATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAAGAAATATCTTACGTGCTCTTCAATATATGTTGGTCTCAGTGCAAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
```

```
QY 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGCCTCTTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGGTTGTAGTATACACAACAGGGATCCTGTACCGAGGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATATAAATGGTCAAAATAGAAATTTAGCATGAGGTGGAGGACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTGACTTTAGATTCGTGTGGCATCCACGGTGCCTCAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTipSerTyr 740
QY 2434 GGAGGATPACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTCCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCAACCTCACCAGATGNAACAGGCTATTACTTAGGATCTGTGGCCATCAAGCAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCCTCTCAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTCACATACCACTATATTAATCTAGTGTCTTTTAGTGAGGCTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGGAGAGACACAGCATAAAGAGTTCTTGAATCCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATCGATCTTTTGCACTACCTTCAAGAAACCTTGGATTCAGTATTTGCTGCTCTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 CTGATA 2859
Db 881 ValIle 882
RESULT 6
US-10-415-122-6
; Sequence 6, Application US/10415122
; Publication No. US2004005369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FFI5217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-122-6
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 4 Gaps: 0
```

US-10-825-632-2 (1-3120) x US-10-415-122-6 (1-882)		
QY	214	ATGGCAGCAGCAATATGGAACAGAACAGCTGGGTGTGAGATATTGAAACTGCCGACTGT 273
Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY	274	GAGGAGAAATATTGAATCACAGGATCGCCTCAATATGGAGCCCTTTTATATGTTGAGCGGTAT 333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY	334	TCCTGGAGTCAGCTTAAAGAGCTTCCTGCGATACAGAAATATATCATGGCTACATGATG 393
Db	41	SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY	394	GCTAAGGCACACATGATTTTCATGTTGTGAAGAGAAATGATCCGATGGACCTCATCA 453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY	454	GACAGATCTATTACCTTGCCATGTCTGTGTGAGAACAGAGAAATACACTGTTTATTCT 513
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY	514	GAATTCCTCCAAACTATCAATAGAGCAGCTCTTAATGCTCTCTTGGAAAGCCTCTTTTG 573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY	574	GATCTTTTTCAGGCAACTGGAATATGGAATGTATTCTCGAGAAGAAGAACTATTAAAG 633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY	634	GAAGAAAAAGCATTCGAAACAGTCGGAATTGCTTCTACGATTATCAACGAAGAGTGA 693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY	694	ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATACGTAACGTAAGAGTGGAGGCCACAGGA 753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY	754	TTTACGCCAACAACTTTAAGGCCCACTCTAGTGGAACTAGTCTCCCAACATACCGATG 813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY	814	GATCCAAATATTATGCCCGCTGATTCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY	874	TGGATATCTAACATCGTAACCAAGAGAAAGAGAGACTCACTTATGTGCAATGAGCTA 933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
QY	934	GCCAAATGGAAGAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA 993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY	994	TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCGAGTGGTGT 1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY	1054	AAAATCTTAGAATTCATATGAAGAAATGATCAATCTGAGTGGAAATTTATCATGTT 1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY	1114	ACATCCCTTATGTTGAAACAAAGAGGAGGCAGATTTCATTCGGTTATCTCTAAAACAGGTACA 1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY	1174	GCNAATCCTTAAGTACATTTTAAGATGTCAGAAATATGATGCTGACGAGGAGGATC 1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluArgIle 340
QY	1234	ATAGATGTCATAGATGAAGCACTAATTCACCTTTGAGATTTCTATTGAGGAGTTGAA 1293
Db		
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY	1294	TATATTGCCAGACTCGATCGACTCTCTGAGGGAATAATATGCTTGCTCATCTACTAGAT 1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY	1354	CGCTCCACAGACTCGCTACAGATAGTGTGTGATCTCACCTGAAATATTATTCACAGTAGAA 1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY	1414	GATGATGTTATGAAAAGGAGAGACTCATTTAGTTCAGTCCCTGATTTCTGTGACCCACTA 1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY	1474	ATTATCTATCAAGAAAAACAACAGACTCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY	1534	CCCCAAAGTCACGAAGAGAAATTTGAGTTTATTTTGGCTCTCGAATGCAAAACAGGTTTC 1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY	1594	CGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT 1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY	1654	GGGCTGCTCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATATACCAGT 1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY	1714	GGTGMAATGGAGTCTTCGCGGCATCGATCTATATATCCAAAGTGTGATGAAGTCAGAGG 1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY	1774	CTGGTATATTTTGAAGGCCACCAAGCTGACTGACCTCCCTTTAGAGCATCACTGTACGTAGTCAGT 1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY	1834	TAGCTAAATCTCGAGAGGTGACAAGCTGACTGACCTGGCTACTCACATCTCTGTGCTGC 1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY	1894	ATCAGTCAGCACTGTGACTTCTTTTATAAGTATAGTAAACCAAGAGAAATCCACACTGT 1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY	1954	GTCTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAAACAAAGGAATTT 2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY	2014	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY	2074	TCTTTTGAAGTACTACTGGATTTCATTTGATGGATGCTCTACAGATGCTCTACAAGCCTCATGATCTA 2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY	2134	CAGCCTGGAAGAAATATCCTACTGTGCTGTTTATATATATGTTGCTCTCAGGTGCAGTTG 2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY	2194	GTGAATAATCCGTTTAAAGGAGTCAGTATTTCCGCTTGATACCTTAGCTCTTAGGT 2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY	2254	TATGCGTGTGATGATAGACACAGGGATTCCTGTCCACGAGGGCTTAAATTTGAAGGC 2313
Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY	2314	GCCTTTAAATATAAAATGGGTCAAATAGAAATTCAGATTCAGGTGGAGGAGCTCCAATAT 2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720

QY 2374 CTAGCTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTAT 2433  
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740  
QY 2434 GAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTGCAGATATCTTCAGGGTTGCTATT 2493  
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760  
QY 2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACTTATATG 2553  
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780  
QY 2554 GGTCACTCCACAGATGACAGGGCTATTACTTAGCATCTGTGGCCATCAGACAGAA 2613  
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800  
QY 2614 AAGTTCCCTCTGAACCAATCGTTTACTGCTTACATGTTTCTCGATGAGAATGTC 2673  
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820  
QY 2674 CATTTTGCACATACCACTATATTAATTAATGCTTTTGTAGTCAGGGCTGGAAGCCCATATGAT 2733  
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840  
QY 2734 TTACAGATCTATCTCCAGGAGACACAGACATAAGAGTTCTCGAATCGGGAGAACATTAT 2793  
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860  
QY 2794 GAATCGATCTTTGCTACCTTCAAGAAACCTTGGATACGATTGCTGCTCTAAAA 2853  
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880  
QY 2854 GTGATA 2859  
Db 881 ValIle 882

## RESULT 7

US-10-825-632-1 ;

; Sequence 1, Application US/10825632

; Publication No. US20040191826A1

; GENERAL INFORMATION:

; APPLICANT: ABBOTT, Catherine Anne

; APPLICANT: GORRELL, Mark Douglas

; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES

; FILE REFERENCE: FCSB-100-Div. 1

; CURRENT APPLICATION NUMBER: US/10/825,632

; CURRENT FILING DATE: 2004-04-15

; PRIOR APPLICATION NUMBER: US 10/070,464

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: PCT/AU00/01085

; PRIOR FILING DATE: 2000-09-11

; PRIOR APPLICATION NUMBER: AU P05709

; PRIOR FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: AU P02762

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 882

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-825-632-1

## Alignment Score#:

Pred. No.:	0	Length:	882
Score:	882.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best local Similarity:	100.0%	Mismatches:	0
Query Match:	85.6%	Indels:	0
DB:	4	Gaps:	0

US-10-825-632-2 (1-3120) x US-10-825-632-1 (1-882)

QY 214 ATGCAGCAGCAATGGAAAAACAGAACAGCTGGGTGTGAGATATTTTAAACCTCCGACTGT 273  
Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGAGAAATTTGAATCACAGATCGGCCTAAATTTGGAGCCCTTTTATTTGTTGAGCGTAT 333  
Db 21 GluGluAsnIleGluSerGlnAspArgPheLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCTTGGAGTCAGCTTTAAAAAGCTGCTGGCATACAGAAATATATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 GCTAAGCAGCACCATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 453  
Db 61 AlaLysAlaProHisaspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 GACAGAAATCTATTACCTTGGCATGCTGGTGTGAACACAGAGAAATATACACTGTTTATTCT 513  
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
QY 514 GAAATTTCCCAAACTATCAATAGACGACGAGCTTTAATGCTCTCTTGAAGCCTCTTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120  
QY 574 GATCTTTTTCAGGCAACACCTGGACTATGGAATGTATTCTCGAGAGAGAACTATTAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140  
QY 634 GAAAGAAAACGCAATGGAAACAGCTCGGAATGCTTCTTACGATTATCACCAAGGAGTGA 693  
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATTCAGTAAAGATGGAGGGCCACAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
QY 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTGTGCCAACATACGAGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCCAAAAATTTATGCCCCGCTGATCCAGACTCGGATGCTTTTATATACATAGCAACGATAT 873  
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTAACATCGTAAACCAGAGAAAGAGAGACTCATTATGTGCACAATGAGCTA 933  
Db 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240  
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAAGAA 993  
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260  
QY 994 TTTGATAGATATTCTGCTGCTATGCTGTGTGTCACAAAGCTGAAACAACTCCCAAGTGTGCT 1053  
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATTTCTTGAATTTCTATATGAAGAAATGTATGAATCTGAGGTGGAATATTATTCATGTT 1113  
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300  
QY 1114 ACATCCCTCTATGTTGAAAAACAGAGGGGAGATTTCAATCCGTTATCTCTAAACACAGGTACA 1173  
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAAACTCTAAAGTCATCTTTTAAAGATGTGAGAAATATGATTCATTCGTAAGAGGATC 1233  
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGTAAGGAACCTAAATCAACCTTTTGAGATTTCTATTTTGAAGAGTTGAA 1293  
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360

QY	1294	TATATTGCCAGCTGATGACTCCTGAGGGAATATGCTTGCTCCATCCTACTAGAT	1353	Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Db	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaIleThrSerIleLeuLeuAsp	380	QY	2434	GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGTCCAGATATCTTCAGGGTTGCTATT	2493
QY	1354	CGCTCCGAGACTCCCTACAGATAGTGTGTGATCTACCTGAATATTATTCACCAAGTAGAA	1413	Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	QY	2494	GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGTATACAGGATACACAGAACTGTATATG	2553
QY	1414	GATGATGTTATGAAAGCGAGAGACTCATTTGAGTCAGTCCCTGATCTCTGTGACCCACTA	1473	Db	761	AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	QY	2554	GGTCACCTCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAACAGAA	2613
QY	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTT	1533	Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	421	IleIleTyrGluGluThrThrAspIleThrIleAsnIleHisAspIlePheHisValPhe	440	QY	2614	AAGTTCCCTCTGAAACCAATCGTTTACTCTTACATGGTTTCTCGATGAGAAATGTC	2673
QY	1534	CCCCAAGTCAACAGAGGAAATTTGAGTTTATTTTTCCTCTGAAATGCAAAACAGGTTTC	1593	Db	801	LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	QY	2674	CATTTTGCACATACCACTGATATTACTAGTCTTTTCTAGTGGCTGGAAAGCCATATGAT	2733
QY	1594	CGTCATTTTACAAAATTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGT	1653	Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	QY	2734	TTACAGATCTATCCTCCAGGAGACACACACATAAGAGTCTCTGAATCCGGAGAACATTAT	2793
QY	1654	GGGCTGCTCTCCCAAGTATTTCAAGTGTCTCTATCAAGAGGATAGCAATTAACAGT	1713	Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500	QY	2794	GAATCGATCTTTTGACCTACCTTCAAGAAAACCTTGGATCACGATTGCTGCTCTAAAA	2853
QY	1714	GGTCAATGGGAAGTCTTGCGCGGCATGATCTATATCCAGTTGATGAAGTCAGAGG	1773	Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Db	501	GlyGluTyrProGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	QY	2854	GTGATA 2859	
QY	1774	CTGCTATATTTGAGGCCAACAAAGACTCCCTTTAGAGCATCACTGTAGTACGTACGT	1833	Db	881	ValIle 882	
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540	RESULT 8			
QY	1834	TACGTAAATCTCGAGAGGTGACAGGCTGACTGACCGGTGCTACTCACATTTCTGTGTC	1893	US-10-982-512-1			
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	; Sequence 1, Application US/10982512			
QY	1894	ATCAGTCAGCACTGTGACTCTTTTATAAGTAAAGTATAGTAAACCAAGAAGATCCACACTGT	1953	; Publication No. US2005059081A1			
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580	; GENERAL INFORMATION:			
QY	1954	GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGGAATTT	2013	; APPLICANT: Qi, Steve			
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	; APPLICANT: Akinsanya, Karen			
QY	2014	TGGGCCACCATTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC	2073	; APPLICANT: Riviere, Pierre			
Db	601	TyrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	; APPLICANT: Junien, Jean-Louis			
QY	2074	TCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA	2133	; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	; FILE REFERENCE: 70669			
QY	2134	CAGCCTGGAAGAAATATCTACTGTGTGTTCATATATGCTGCTCTCAGGTGCAGTTG	2193	; CURRENT APPLICATION NUMBER: US/10/982,512			
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyLysProGlnValGlnLeu	660	; CURRENT FILING DATE: 2004-11-05			
QY	2194	GTGAATATCGTTTAAAGAGTCAAGTATTTCCCTTGAATACCTAGCTCTCTAGGT	2253	; PRIOR APPLICATION NUMBER: US/09/976,674			
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680	; PRIOR FILING DATE: 2001-10-12			
QY	2254	TATGTGTTGTAGTATAGACAAACAGGGATCTCTGTCCACGAGGGCTTAAATTTGAAGGC	2313	; PRIOR APPLICATION NUMBER: US 60/240,117			
Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700	; PRIOR FILING DATE: 2000-10-12			
QY	2314	GCCTTTAAATATATAATGGTCAATATAGAAATTTGACGATCAGGTGGAGGAAGCTCCAATAT	2373	; NUMBER OF SEQ ID NOS: 61			
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720	; SOFTWARE: PatentIn version 3.1			
QY	2374	CTAGCTTCTCGATATGATTTTCACTTAGACTTAGATCGTGTGGGCATCCACGGCTGGCTCAT	2433	; SEQ ID NO 1			
				; LENGTH: 882			
				; TYPE: PRT			
				; ORGANISM: Homo sapiens			
				US-10-982-512-1			
				Alignment Scores:			
				Pred. No.:	0	Length:	882
				Score:	882.00	Matches:	882
				Percent Similarity:	100.0%	Conservative:	0
				Best Local Similarity:	100.0%	Mismatches:	0
				Query Match:	85.6%	Indels:	0
				DB:	5	Gaps:	0
				US-10-825-632-2 (1-3120) x US-10-982-512-1 (1-882)			
				QY 214 ATGGCAGCAGCAATGAAACAGAACAGCTGGGTGTGTGAGATATTTGAAACTGCCGACTGT	273		
				Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20			

QY	274	GAGGAGAAATTGAATCACAGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGGGTAT	333	QY	1354	CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCCTGAANTTATTATCCAGTAGAA	1413
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40	Db	381	ArgSerGlnThrArgLeuGlnIleValIleuIleSerProGluLeuPheIleProValGlu	400
QY	334	TCTCGAGTCACGCTTAAAAAGCTGCTCCGATACACAGAAAATATCATGGCTACATGATG	393	QY	1414	GATGATGTTATGGAAGCAGAGACTCAATCAGTCAGTCGCTGATCTCTGTGAGCCACTA	1473
Db	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60	Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	394	GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATCA	453	QY	1474	ATTATCTATGAAGAAACAACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTT	1533
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80	Db	421	IleIleTyrGluGluThrThrAspIleThrPheAsnIleHisAspIlePheHisValPhe	440
QY	454	GACAGAAATCTATTACCTTCCCATGCTGCTGAGACAGAGAAATACACTGTTTTATTCT	513	QY	1534	CCCCAAGTCACGAAGAGAAATAGAGTTTATTTTGGCTCTGATCCAAACAGGTTTC	1593
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100	Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	514	GAATTTCCCAAACTATCAATAGACGACGAGCTTAATGCTCTCTGGAGCCCTCTTTTG	573	QY	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGAAAGCAATAATAACCATCCAGTGGT	1653
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120	Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	574	GATCTTTTTCAGGCAACACTGGAATGATGTAATGTTCTCGAGAGAGAACTATTAGA	633	QY	1654	GGCTCGCTCGCTCCAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACAGT	1713
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140	Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	634	GAAGAAGACGATTTGGACAGCTCGGAATTCCTTACGATTTATCACCAAGCAAGTGA	693	QY	1714	GGTGAATGGGAAGTTCTTTGGCCGCGATCGATCTAATATCCAAGTTGATGAAGTCAGAGG	1773
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGAGGGCCACAAAG	753	QY	1774	CTGGTATATTTGAAGGCACCAAGACTGCCCTTTAGAGCATCACCTGTAGTAGTCAGT	1833
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
QY	754	TTTACGCAACACTTTAAGGCCAATCTAGTGGAAACTAGTTGTCACACATACCGATG	813	QY	1834	TAGCTAAATCCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1893
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	814	GATCCAAATATTGCCCCCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT	873	QY	1894	ATCAGTCAGCACTGTGACTCTTTTATAAGTAAATAGTAAACCAAGAAATCCACACTGT	1953
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	874	TGGATATCTAACATCTAAACAGAGAGAAAGAGACTCACTTATGTGCACAAATGAGCTA	933	QY	1954	GTGTCCCTTTACAGCTATCAAGTCTCAAGTCTGAGATGACCAACTGCACAAACAAAGAAATTT	2013
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	934	GCCAAACATGAAGAGATGCCAGATCAGCTGGAGTCGCTACTCTTGTCTCCAAAGAGAA	993	QY	2014	TGGGCCACCAATTTTGATTCAGCAGCTCCTCTCTCTGACTATCTCTCCAGAAATTTTC	2073
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	994	TTTGATAGATTTCTGGCTATTGGTGTGTCCAAAAGCTGAAACAACTCCCAAGTGGGT	1053	QY	2074	TCTTTTCAAAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCCTCATGATCTA	2133
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrProSerGlyGly	280	Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	1054	AAAACTCTAGATTTCTATAGAATAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113	QY	2134	CAGCTCGGAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCCAGTCAGTTG	2193
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	1114	ACATCCCTCATTTGTGAAAACAAGGAGGCGAGATTTCATTCCGTTATTCCTAAACAGGTACA	1173	QY	2194	GTGAAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGCCCTCTAGGT	2253
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	1174	GCAATCCTAAAGTCATTTTAAGATGTCAGAAATAATGATGATGCTGAGGAGGATC	1233	QY	2254	TATGTGTTGTAGTAGTAGACAAACAGGGATCCTGTCCCGAGGCTCTAAATTTGAAGGC	2313
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
QY	1234	ATGATCTCATAGTAGGAACCTAATTCACCTTTTGGATCTCTATTTTGAAGGAGTTGAA	1293	QY	2314	GCCTTTAAATATAAATGGGTCAAAATAGAAATTCAGTACGCTGGAAGGACTCCAATAT	2373
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	1294	TATATTGCCAGCTGGAATCGACTCTCGAGGGAATAATATGCTTGGTCCATCCTACTAGAT	1353	QY	2374	CTAGCTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTAT	2433
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
				QY	2434	GGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTCAGATATCTTTCAGGGTTGCTATT	2493

```
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCGAGTCACTCTGTCGATCTTCTATGATACAGGATACACGGACGCTATATG 2553
Db 761 AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyThrGluArgTyrMet 780
QY 2554 GGTCACTCAGCAGATGAGCAGGCTATTACTTACGATCTTACAGTGTCTCTGATGAGATGTC 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCCTCTGAAACCAATCGTTTACGCTCTTACATGAGTTCCTCGATGAGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTCCACATACCGATATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGGAGACACAGCATAGAGTTCTCGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATCGATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGATTTGCTGCTCTAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 9
US-09-976-674-19
; Sequence 19, Application US/09976674
; Patent No. US2020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19

Alignment Scores:
Pred. No.: 0 Length: 658
Score: 655.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.6% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-19 (1-658)
QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTCGGACTGT 273
Db 1 MetaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATCACAGGATCGGCCTAAATTTGAGCGCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGTTGCCGATACCCAGAAATATCATGGCTACATGATG 393
```

```
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAC 453
Db 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAAATATTACCTTGCATGCTCGTGTGAGAACAGAGAAAAATACACTGTTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCACACTGGACTATGGAATGTATTCGAGAGAGAAAGAACTATTAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCAATTCGAAACAGTCGGAATTCCTTACGATTATACACAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCAGTAAAGATGAGAGGCCCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATTGCCCGCTGATCCAGACTGGATGTGCTTTTATACATAGCAACGATTT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAAACAGAGAAAGAGACTCACTTATGTGCACAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATCGAAAGAGATGCCAGATTCAGCTGGAAGTCGTACCTTTGTTCCTCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCCACTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTTAGAATTTCTATATGAAGAAATGATGAATCTCAGGTGGAAATATTTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTTATGTTGAAAACAGGAGGCGAGATTCATTCCGTTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCTTAAAGTCACATTTTAAAGATGTGAGAAATATATGCTTGGTCCATCTCTAGAT 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACTAATTCACACCTTTTGAGATTCATATTCAGAGAGTTGA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGCTGGATGGAATCTGCTCAGGAAATATGCTTGGTCCATCTCTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCGAGCTCGGCTACAGATAGTGTGATCTCACCTGAATATTATTCACAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTTATGGAAGGCGAGACTCATTTGAGTCTAGTGCCTGATTTCTGTGACGCCACTA 1473
```







```
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgGlyProLysThrGlyThr 320
QY 1174 GCAATCTTAAGTCACCTTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGAGATGCTATTGAAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATGCCAGAGCTGGATCGACTCCTCGAGGAAATATGCTTGCTTGCCTACTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGAGCTCGCCCTACAGATAGTGTTCATCTCACCTGAAATATTATCCACAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGGCAGAGACTCATTGAGTCAGTGCCTGATTCCTGCACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACAACAGACATCTGATAAATATCCATGACATCTTTTCATGTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCACGAGGAATGAGTTTATTTTGGCTCTGATGCAACCAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATAACGATCCAGTCGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGCCTGCTGCTCCAAAGTATTCAAAGTGTCTCTATCAAGAGAGAGATAGCAATACCACT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTGTGCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGCTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTCCTGTGTC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTACAGCTATCAAGTCTCTGAAGATGCCCAACTTGCARAAACAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluLysAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGCGCCACCAATTTGGATTTCAGAGTCTCTCTCTGACTATCTACCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655
```

RESULT 11

US-09-976-674-11

```
; Sequence 11, Application US/09976674
; Patent No. US20020115943A1
; GENERAL INFORMATION:
; APPLICANT: Q1, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-11

Alignment Scores:
Pred. No.: 0 Length: 661
Score: 655.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Beat Local Similarity: 100.0% Mismatches: 0
Query Match: 63.6% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-11 (1-661)

QY 214 ATGCACAGCAATGGAAACAGACAGCTGGTGTGAGATATTTGAAACTCCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATGAATCACAGATCGGCCCTAAATTGGACGCTTTTATGTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTGCCGATACCAGAAAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGCACCACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAAATCTATTACCTTGCCTGCTGATGAGCAGAGCTTAAATGCTCTCTTGAAGCTCTTTTG 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAACTATCAATAGAGCAGAGCTTAAATGCTCTCTTGAAGCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysAspProLeuLeu 120
QY 574 GATCTTTTTCAGGCACACCTGGACTATGGAATGTAATCTTCGAGAGAAGAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAACGCATTCGAAACAGTCGAAATTCCTTACGATTATCAGTAAGAGAGGCCACAGGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAGCCCGTAGTGAATTTATCAGTAAGAGAGAGAGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTTACGCAACACCTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGAGATG 813
Db 181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAAATATGCCCCGCTGATCCAGACTGGATGCTTTTATATACATAGCAACGATATT 873
```



101 GluileProLysThrIleAsnArgAlaIleValLeuMetLeuSerTrpIysProLeuLeu 120  
120 |||||  
574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATCTTCGAGAGAGAACTATTAGA 633  
633 |||||  
121 AspLeuPheGlnAlaThrLeuAspTrpGlyMetCysSerArgGluGluGluLeuArg 140  
140 |||||  
634 GAAAGAAAACGCATTTGAAACAGTCGGAATGCTTCTTACGATATTCACCAAGGAAGTGA 693  
693 |||||  
141 GluArgLysArgIleGlyThrValGlyIleAlaSerTrpAspTrpHisGlnGlySerGly 160  
160 |||||  
694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATACGTAAAGATGGAGGCCCAACAGGA 753  
753 |||||  
161 ThrPheLeuPheGlnAlaGlySerGlyIleTrpHisValLysAspGlyGlyProGlnGly 180  
180 |||||  
754 TTTACGCCAACACCTTTAAGCCCAATCTAGTGGAACTAGTGTGTCACCAATACGGATG 813  
813 |||||  
181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
200 |||||  
814 GATCCAAAATATTGCCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873  
873 |||||  
201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
220 |||||  
874 TGGATATCTAACATCGTAACACAGAGAAAGAGACTCACTTATGTGCACAATGAGCTA 933  
933 |||||  
221 TrpIleSerAsnIleValThrArgGluGluArgAlaGlyValAlaThrPheValHisAsnGluLeu 240  
240 |||||  
934 GCCAACATCGAAGAAGATGCCAGATCAGCTGGAGTCGCTACTTGTGTTCTCCAGAGAGAA 993  
993 |||||  
241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValHisAsnGluLeu 260  
260 |||||  
994 TTTGATAGATATTCTCGCTATTGTTGGTGTCCBAAAGCTGAAACNACTCCCAAGTGGGT 1053  
1053 |||||  
261 PheAspArgTrpSerGlyTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
280 |||||  
1054 ARAATCTTGAATTTCTATACAGAAATGATGATCTGAGGTGGAAATTTATCATGTT 1113  
1113 |||||  
281 LysIleLeuArgIleLeuTrpGluAsnAspGluSerGluValGluIleHisVal 300  
300 |||||  
1114 ACATCCCTATGTTGAAACACAGAGGCGCAGATTCATTCGTTATCTCTAAACAGGTPACA 1173  
1173 |||||  
301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTrpProLysThrGlyThr 320  
320 |||||  
1174 GCAATCCTAAAGTCACTTTTAAAGTGTGAGAAATATGATGATGCTGGAAGGAGATC 1233  
1233 |||||  
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
340 |||||  
1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTTCTATTGAGGAGTTGAA 1293  
1293 |||||  
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
360 |||||  
1294 TATATTGCCAGAGCTGGAGTCTCTCGAGGAAATATGCTTGGTCCATCTCTACTAGAT 1353  
1353 |||||  
361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTrpAlaTrpSerIleLeuLeuAsp 380  
380 |||||  
1354 CGCTCCCACTCGCTCAGAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA 1413  
1413 |||||  
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
400 |||||  
1414 GATGATGTTATGAAAGGAGAGACTCATTGAGTCAGTGCCTGATCTGTGAGCGCCACTA 1473  
1473 |||||  
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
420 |||||  
1474 ATTATCTATGAGAAACACACATCTGGATTAATATCCATGACATCTTCTGCTTTT 1533  
1533 |||||  
421 IleIleTrpGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
440 |||||  
1534 CCCCAGGTCAGAGAGGAAATTCAGTTATTTTTCCTCTGATGCAATGCAAAACAGTTTC 1593  
1593 |||||  
441 ProLysSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
460 |||||  
1594 CGTCACTTTATACAAATTCATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGT 1653  
1653 |||||

461 ArgHisLeuTrpLysIleThrSerIleLeuLysGluSerLysTrpLysArgSerSerGly 480  
480 |||||  
1654 GGGCTGCTGCTCCAGTGTCAAGTGTCAAGTGTCTCTATCAAAAGAGAGATAGCAATACCACT 1713  
1713 |||||  
481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500  
500 |||||  
1714 GGTGAATGGGAAGTCTTTCGCCGCGCATGGATCTAATATATCAAAAGTGTGAGAAAGTGAAGG 1773  
1773 |||||  
501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520  
520 |||||  
1774 CTGTATATATTTTGAAGGCACAAAGACTCCCTTTTAGAGCATCACTCTAGTGTAGTCACT 1833  
1833 |||||  
521 LeuValTrpPheGluGlyThrLysAspSerProLeuGluHisHisLeuTrpValValSer 540  
540 |||||  
1834 TACATAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCATCTTTCTGTGTCG 1893  
1893 |||||  
541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTrpSerHisSerCysCys 560  
560 |||||  
1894 ATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953  
1953 |||||  
561 IleSerGlnHisCysAspPhePheIleSerLysTrpSerAsnGlnLysAsnProHisCys 580  
580 |||||  
1954 GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGTACCCAACTTCGCAAAACAAGGAATTT 2013  
2013 |||||  
581 ValSerLeuTrpLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
600 |||||  
2014 TGGCCACCACTTTTGGATTTCAGAGGTCTCTCTGCTGATATCTCTCCGAAATTTTC 2073  
2073 |||||  
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTrpThrProGluIlePhe 620  
620 |||||  
2074 TCTTTTGAAGTACTACTGATTTACATTTGATTTGAGGATGCTCTACAGCCTCATGATCTA 2133  
2133 |||||  
621 SerPheGluSerThrThrGlyPheThrLeuTrpGlyMetLeuTrpLysProHisAspLeu 640  
640 |||||  
2134 CAGCTCGAAGAAATATCTCTGCTGTCTTCAATATATGTTGGT 2178  
2178 |||||  
641 GlnProGlyLysLysTrpProThrValLeuPheIleTrpGlyGly 655  
655 |||||  
RESULT 13  
US-09-976-674-7  
; Sequence 7, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-7  
Alignment Scores:  
Pred. No.: 0 Length: 690  
Score: 655.00 Matches: 655  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 63.6% Indels: 0  
DB: 3 Gaps: 0  
US-10-825-632-2 (1-3120) x US-09-976-674-7 (1-690)  
Qy 214 ATGGCAGCAGCAATGAAACAGACAGCTGGTGTGAGATATTTGAACTCGGACTGT 273  
273 |||||

Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluGluPheGluThrAlaAspCys 20  
Qy 274 GAGAGAAATATGAATCACAGATCGGCTAAATGGAGCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
Qy 334 TCCTGGAGTCAGCTTAAAGAGCTGCTGCCATACCAAGAAATATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
Qy 394 GCTAAGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 453  
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
Qy 454 GACAGATCTATTACCTGCCATGCTGGTGAGAACAGAGAAATACACTGTTTATTCT 513  
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100  
Qy 514 GAAATCCCAAACTATCAATAGACGACAGTCTTAATGCTCTCTCGAAGCCTCTTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
Qy 574 GATCTTTTTCAGCAACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140  
Qy 634 GAAAGAAACGCAATTCGAACAGTCGGAATGCTCTTACGATTATCAACGAAGAGTGA 753  
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
Qy 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATACGTAAAAAGATGGAGGCCCAAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180  
Qy 754 TTACGCAACAACTTTAAGGCCAATCTAGTGGAAACTAGTTGTCCTCCACATACGGATG 813  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
Qy 814 GATCCAAATATGCCCCCTGATCCAGACTGATGCTTTTATACATAGCAACGATATT 873  
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
Qy 874 TGAATATCTTAACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCACATGACCTA 933  
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240  
Qy 934 GCCAATGGAAGAAGATCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAGAGAGAA 993  
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260  
Qy 994 TTTGATAGATATTCTGGCTATTGCTGTGTCCTCAAAAGCTGAAACAACTCCCGAGTGGT 1053  
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
Qy 1054 AAAATTTCTAGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAAATTTATCATGTT 1113  
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300  
Qy 1114 ACATCCCTATGTTGGAAAACAAGAGGCGAGATTCTCCGTTATCTTAAACAGGTACA 1173  
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
Qy 1174 GCAATCTTAAGTCATTTTAAGATGTCAGAAATAATGATGATGCTGGAAGAGGATC 1233  
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
Qy 1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTTCAGATTTCTATTGAAAGGAGTTCAA 1293  
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360  
Qy 1294 TATATTGCCAGAGCTGGATGACTCTCGAGGGAATAATGCTGGTCCATCTCTACTAGAT 1353  
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380

Qy 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413  
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
Qy 1414 GATGATGTTATGAAAGGAGAGACTCATTTAGTCAGTCGCTGCTGTGTCAGCCACTA 1473  
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
Qy 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533  
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
Qy 1534 CCCAAAGTCACGAAGAGAAATTCAGTTTATTTTTCCTCTGATTCGAAACAGGTTTC 1593  
Db 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
Qy 1594 CGTCATTTATACAAATTCATCTATTAAAGGAAGCAATATAAACCATCCAGTGGT 1653  
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480  
Qy 1654 GGGCTGCTCTCCAAAGTATTTCAAGTGTCTCTATCAAGAGAGATAGCAATTTACAGT 1713  
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500  
Qy 1714 GGTCAATGGGAAGTCTTGGCCGATGGATCTAATATCTCAAGTTGATGAAGTCAGAGG 1773  
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520  
Qy 1774 CTGGTATATTTGAAGGCCAACAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833  
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540  
Qy 1834 TACGTAATCTCGAGAGGTGACAAGGCTGACGACCGTGGCTACTCACATTTCTGCTGC 1893  
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560  
Qy 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTATAGTATAGTAACCAAGAAATCCACACTGT 1953  
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580  
Qy 1954 GTGTCCTTTTACAGCTATCAAGTCTCTGAGATGACCAACTTGCACAAACAAAGAAATTT 2013  
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600  
Qy 2014 TGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTCTGACTATCTCTCCAGAAATTTTC 2073  
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620  
Qy 2074 TCTTTTGAAGTACTACTGGATTTTACATGTTATGGATGCTCTACAGCCTCATGATCTA 2133  
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640  
Qy 2134 CAGCCTCGAAGAAATATCTACTGCTGCTTCATATATGTTGGT 2178  
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

## RESULT 14

US-10-982-512-7  
; Sequence 7, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; CURRENT FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117

```
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-7

Alignment Scores:
Pred. No.: 0          Length: 690
Score: 655.00        Matches: 655
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%    Mismatches: 0
Query Match: 63.6%             Indels: 0
DB: 5                         Gaps: 0

US-10-825-632-2 (1-3120) x US-10-982-512-7 (1-690)

QY 214 ATGCACAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTGAAACTGGCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

QY 274 GAGGAGAAATTTGAATCAGAGATCGGCTAAATTTGGAGCCCTTTTATGTTGGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluAspTyr 40

QY 334 TCCTGGAGTCAGCTTAAAAAGCTGTTCGCCGATACAGAAAAATATCATGGCTCATGATG 393
Db 41 SerTTPSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80

QY 454 GACAGAACTATTACCTCCCATGCTCTGTCGACGACGAGAAATACACTTTTATTTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100

QY 514 GAAATTCCTTTCAGGCACTGACATGGAATGATTCCTGAGAAAGAACTATTAAAG 633
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120

QY 574 GATCTTTTTCAGGCACTGACATGGAATGATTCCTGAGAAAGAACTATTAAAG 693
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140

QY 634 GAAAGAAACGCTTGGAAACAGTCGGAATTCCTTACGATTATACCAAGGAAGTGA 693
Db 141 GluArgGlyArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAAAGATGGAGGCCAC 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

QY 754 TTTTACGCAACAACTTTAAGGCCCAATCTAGTGGAATCTAGTTGTCCCAACATACG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGAT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220

QY 874 TGGATATCTAATCATCGTAACAGAGAGAGGAGCTCACTTATGTGACATAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgGluThrTyrValHisAsnGluLeu 240

QY 934 GCCAATCGGAGAGAGAGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260

QY 994 TTTGTAGATATTCTGGCTATTGTTGTTGTTCCAAAAGCTGAAACAACTCCCAAGTGGT 1053
Db 1053
```

```
Db 261 PheAspArgTyrSerGlyTyrTrrTrrPrrCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTTGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAAATATTTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTTATGTTGGAAACAAGAGGGCAGATTTCATCCGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCCTAAAGTCATCTTTTAAGATGTGCAGAAAATAATGATTGATGCTGAAAGGAAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACTAATTCACCTTTTGGAGATTCCTATTTTGAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCACAGAGCTGGATGGACTCTCTGAGGAAAAATATGCTTGGTCCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrrThrProGluGlyLysTyrAlaTrrPrrSerIleLeuLeuAsp 380
QY 1354 CGCTCCACAGACTCGCTCAGAGAGACTCATTCAGCTGAGTCCCTGATTCGTGACCCACTA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAGGCAGAGACTCATTCAGCTGAGTCCCTGATTCGTGACCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAAGAAACAACAGACATCTGGATATAATATCCATGACATCTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCACAGAGAGAAATTCAGTTTATTTTTCCTCTGAATCAAAACAGGTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTACATCTATTTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGCTCCCTCGCTCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAAAATTTACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TAGCTAAATCTGGAGAGGTGACAAGCTGACTGACCGCTGGCTACTCACAATTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys 560
QY 1894 ATCAGTCAGACTGTGACTTCTTTTATAGTATAGTATAGTAAACAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTTACAGCTATCAAGTCTGAGAGTACCCAACTTGCAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCCACTTTTGGATTTCAGCAGGTCTCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAGGCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
```

QY 2134 CAGCCTGGAAGAAATATCTCTACTGTCTGTTTCATATATGCGTGT 2178  
|||||  
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

RESULT 15

US-09-976-674-21  
; Sequence 21, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; FILE REFERENCE: 70669  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 613  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-21

Alignment Scores:  
Pred. No.: 0 Length: 613  
Score: 607.00 Matches: 607  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 58.9% Indels: 0  
DB: 3 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-21 (1-613)

QY 214 ATGCACGACGAATGGAACAGACAGCTGGGTGTGAGATATTGAACTCGCGACTGT 273  
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20  
QY 274 GAGGAGATATTGAATCACAGGATCGGCTAAATTCGAGCCCTTTTATGTTGAGCGGTAT 333  
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40  
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACCGAATAATCATGGCTACATGATG 393  
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60  
QY 394 GCTAAGGCACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATCA 453  
Db 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80  
QY 454 GACGAGATCTATTACCTTGGCATGCTGTGGAGAACAGAGAAATACACTGTTTATCT 513  
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgLysThrLeuPheTyrSer 100  
QY 514 GAAATCCCAAACTATCAATAGACGACGAGCTTAACTGCTCTCTTGGAGCCCTTTTG 573  
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120  
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTAATTCGAGAGAAGAACTATTAGA 633  
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140  
QY 634 GAAAGAAACCATGGAACAGTCGGAATTCCTTACGATTATCACCAAGAGGTGGA 693  
Db 141 GluArgLysAlaGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160  
QY 694 ACATTTCTGTTTCAAGCGGTAGTGAATTTATCAGTAAAGATGGAGGCCCAAGGA 753  
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

QY 754 TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATAGCGATG 813  
|||||  
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200  
QY 814 GATCCAAAATTTATCCCCGCTGATCCAGACTGAGTGTCTTTTATACATACCAACGATATT 873  
|||||  
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220  
QY 874 TGGATATCTAACATCGTAACCCAGAGAAGAGAGACTCATTATGTGACACATGAGCTA 933  
|||||  
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240  
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTGTCTCCCAAGAGAA 993  
|||||  
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260  
QY 994 TTTGATAGATATTCTGCTATTGGTGTGTCCAAAAGCTGAAAACACTCCCAAGTGTGTGT 1053  
|||||  
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280  
QY 1054 AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTCAGGTGGAAATTTATCATGTT 1113  
|||||  
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300  
QY 1114 ACATCCCTATGTTGAAACCAAGAGGGCAGATTCATTCCTGTTATCCTAAAAACAGGTACA 1173  
|||||  
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320  
QY 1174 GCAATCTCTAAAGTCACCTTTTAGATGTGAGAAATATGATTCATGCTGCAAGGAGGATC 1233  
|||||  
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340  
QY 1234 ATAGATGTCATAGATAAGGAATAATTCACCTTTTGGAGATTCATTTTCAAGAGAGTTGAA 1293  
|||||  
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlyValGlu 360  
QY 1294 TATATTGCGAGCTGGATGACTCTCTGAGGAAAAATATGCTGTGGTCCATCTCTACTAGAT 1353  
|||||  
Db 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380  
QY 1354 CGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACCTGAATTTATTTATCCAGTAGAA 1413  
|||||  
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400  
QY 1414 GATGATGTTATGGAAGGCAGAGACTCATTGAGTCAGTCGCTGATTCGTGACGCCACTA 1473  
|||||  
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420  
QY 1474 ATTATCTATGAAGAAACACAGACATCTGATATAATATCCATGACATCTTTCATGTTTTT 1533  
|||||  
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440  
QY 1534 CCCCAAGTACGGAAGAGGAATTTAGTTTATTTTTCCTCTGAAATGCAAAACAGGTTTC 1593  
|||||  
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460  
QY 1594 CQTCAATTTATACAAATTTACATCTATTATTAAGGAAGCAAAATATATAACGATCCAGTGT 1653  
|||||  
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGly 480  
QY 1654 GGGCTGCTCTCCCAAGTGTATTTCAAGTGTCTTATCAAAAGAGAGATAGCAATTTACCACT 1713  
|||||  
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500  
QY 1714 GGTGAATGGGAAGTCTTTCGCGCGCATGATCTTAATATCAAGTGTGATCAAGTCAAGAG 1773  
|||||  
Db 501 GlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520  
QY 1774 CTGGTATATTTTGAAGGCACAAAGACTCCCTTTTACAGCATCCTGTACGTAGTCACT 1833  
|||||  
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540

QY	1834	TACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTCTTGCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCAGCTGTGACTTCTTTATAGTTAGTATAGTAACGAGAGAAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCCTTTACAGCTATCAAGTCCTGAGATGACCCCAACTTGCAAAACAAAGGAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACCAATTTTGATTCA	2034
Db	601	TrpAlaThrIleLeuAspSer	607

Search completed: May 2, 2006, 05:00:52  
Job time : 505.5 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: May 2, 2006, 04:46:40 ; Search time 11.5 Seconds  
(without alignments)  
2467.668 Million cell updates/sec

Title: US-10-825-632-2  
Perfect score: 1030  
Sequence: 1 aagtgctaaagctccgagg.....agaattactataaaaaaaa 3120

Scoring table: **OLIGO**  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 60.0, Fgapext 7.0  
Delop 60.0, Delext 60.0

Searched: 232119 seqs, 45477862 residues

Word size: 1

Total number of hits satisfying chosen parameters: 424616

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame\_n2p.model -DEV=xlp  
-Q=/abss/ABSSWEB.spool/US10825632/runat\_01052006\_110250\_4729/app\_query.fasta\_1  
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=oligo.n2p.rapbn  
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abs02p  
-USER=US10825632 @CGN 1.1 34 @runat 01052006 110250 4729 -NCPU=6 -ICPU=3  
-NO\_WMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=60 -DELEXT=60

Database : Published Applications AA New:

- 1: /SID55/ptodata/2/pubpaa/US08\_NEW\_PUB.pep1.\*
- 2: /SID55/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /SID55/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /SID55/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /SID55/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 6: /SID55/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 7: /SID55/ptodata/2/pubpaa/US09\_NEW\_PUB.pep1.\*
- 8: /SID55/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 9: /SID55/ptodata/2/pubpaa/US10\_NEW\_PUB.pep1.\*
- 10: /SID55/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 11: /SID55/ptodata/2/pubpaa/US11\_NEW\_PUB.pep1.\*
- 12: /SID55/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	882	85.6	882	11	US-11-151-601-20
2	109	10.6	109	11	Sequence 20, Appl
3	24	2.3	24	11	Sequence 10, Appl
4	11	1.1	99	11	Sequence 16, Appl
5	11	1.1	349	11	Sequence 11, Appl
					Sequence 1591, Ap

6	11	1.1	497	11	US-11-264-096-1593	Sequence 1593, Ap
7	11	1.1	497	11	US-11-264-096-1594	Sequence 1594, Ap
8	9	0.9	353	11	US-11-172-740-719	Sequence 719, App
9	8	0.8	49	11	US-11-071-259-3	Sequence 3, Appli
10	8	0.8	90	11	US-11-096-568A-9319	Sequence 9319, Ap
11	8	0.8	108	9	US-10-504-588-6	Sequence 6, Appli
12	8	0.8	108	11	US-11-073-420-6	Sequence 6, Appli
13	8	0.8	123	11	US-11-194-246-399	Sequence 399, App
14	8	0.8	129	11	US-11-096-568A-14806	Sequence 14806, A
15	8	0.8	129	11	US-11-096-568A-26015	Sequence 26015, A
16	8	0.8	137	11	US-11-087-099-943	Sequence 943, App
17	8	0.8	147	11	US-11-087-099-10322	Sequence 10322, A
18	8	0.8	149	11	US-11-096-568A-26340	Sequence 26340, A
19	8	0.8	151	11	US-11-096-568A-24024	Sequence 24024, A
20	8	0.8	157	11	US-11-096-568A-19057	Sequence 19057, A
21	8	0.8	164	11	US-11-096-568A-14805	Sequence 14805, A
22	8	0.8	164	11	US-11-096-568A-26014	Sequence 26014, A
23	8	0.8	174	11	US-11-072-512-2398	Sequence 2398, Ap
24	8	0.8	177	11	US-11-096-568A-14804	Sequence 14804, A
25	8	0.8	177	11	US-11-096-568A-26013	Sequence 26013, A
26	8	0.8	187	11	US-11-096-568A-24022	Sequence 24022, A
27	8	0.8	200	11	US-11-096-568A-19056	Sequence 19056, A
28	8	0.8	200	11	US-11-096-568A-26339	Sequence 26339, A
29	8	0.8	224	11	US-11-096-568A-15343	Sequence 15343, A
30	8	0.8	233	11	US-11-188-298-2947	Sequence 2947, Ap
31	8	0.8	245	11	US-11-096-568A-10277	Sequence 10277, A
32	8	0.8	246	11	US-11-096-568A-19055	Sequence 19055, A
33	8	0.8	249	11	US-11-096-568A-13087	Sequence 13087, A
34	8	0.8	266	11	US-11-096-568A-19487	Sequence 19487, A
35	8	0.8	270	11	US-11-098-686-10898	Sequence 10898, A
36	8	0.8	281	11	US-11-096-568A-15342	Sequence 15342, A
37	8	0.8	287	11	US-11-080-991-66	Sequence 66, Appl
38	8	0.8	295	11	US-11-096-568A-13086	Sequence 13086, A
39	8	0.8	303	9	US-10-517939-214	Sequence 214, App
40	8	0.8	309	11	US-11-172-740-1121	Sequence 1121, Ap
41	8	0.8	311	11	US-11-188-298-21052	Sequence 21052, A
42	8	0.8	315	11	US-11-096-568A-19486	Sequence 19486, A
43	8	0.8	317	11	US-11-184-298-19438	Sequence 19438, A
44	8	0.8	319	9	US-10-0558977-290	Sequence 290, App
45	8	0.8	323	11	US-11-096-568A-13085	Sequence 13085, A

ALIGNMENTS

RESULT 1

- US-11-151-601-20
- Sequence 20, Application US/1151601
- Publication No: US20060003413A1
- GENERAL INFORMATION:
- APPLICANT: Millennium Pharmaceuticals, Inc.
- APPLICANT: Meyers, Rachel E.
- APPLICANT: Olandt, Peter J.
- APPLICANT: Kapeller-Libermann, Rosana
- APPLICANT: Curtis, Rory A. J.
- APPLICANT: Williamson, Mark
- APPLICANT: Welch, Nadine
- TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.
- TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
- FILE REFERENCE: MP100-054FRCPIOMNIDVIM
- CURRENT APPLICATION NUMBER: US/11/151.601
- CURRENT FILING DATE: 2005-06-13
- PRIOR APPLICATION NUMBER: US 10/170,789
- PRIOR FILING DATE: 2002-06-13
- PRIOR APPLICATION NUMBER: US 09/797,039
- PRIOR FILING DATE: 2001-02-28
- PRIOR APPLICATION NUMBER: PCT/US01/06525
- PRIOR FILING DATE: 2001-02-28
- PRIOR APPLICATION NUMBER: US 60/186,061
- PRIOR FILING DATE: 2000-02-29
- PRIOR APPLICATION NUMBER: US 09/882,166
- PRIOR FILING DATE: 2001-06-15
- PRIOR APPLICATION NUMBER: PCT/US01/19269
- PRIOR FILING DATE: 2001-06-15

```
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 882.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.6% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-151-601-20 (1-882)
QY 214 ATGCGACGCAATGGAACAGACAGCTGGGTGTGAGATTTGAACTGGCGACTGT 273
DB 1 MetAlaAlaMetGluThrGluLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGAGAAATATTGAATCAGAGATCGGCGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTGCGATACAGAAAATATCATGGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGCACCACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTC 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGGCATGCTGTGGTGAGAACAGAGAAAATACACTGTTTATTTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAATCATCAATAGACGACGATGGAATGTTTCTCTTGGAAAGCCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGGCAACACTGGACTATGGAATGTATTCCTCGAGAGAGAACTATTAAAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGRAACGCATTGGAACAGTCGGAAATGCTTACGATTATCACAGGAGAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGTAGTGGAAATTTATCATCGTAAAGATGGAGGCCCAAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTTAAGCCCAATCTAGTGGAAACTAGTTGTGCCAAACATACGGATG 813
DB 181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATATTATGCCCCGTGATCCAGACTGGATGCTTTTATATCATAGCAACGATATT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAACAGAGAGAAAGAGACTCATCTTATGTGCCAAATGAGCTA 933
```

```
DB 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAGAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGCTGTCCAAAGCTGAAACCAACATCCAGTGGTGGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTCTTACAATTTATATGAAGAAAATGATGAATCTGAGTGGAAATATTATCATCTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTCATGTTGAAACACAGAGGCGAGATTTCATCCGTTATCCCTAAACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCTTAAAGTCACCTTTTAAGATGTCAGAAAATTAATGATTGATGCTGAAAGGAGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAATTCACACCTTTTGGAGATTCTATTTGAAGGAGTTCAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCCAGACGTGATGGACTCCTGAGGAGAAATATGCTTGGTCCATCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCCTGAATTTATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAGGACAGAGACTCATTTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAAGTCACAAAGAGAAATTTGAGTTTATTTTGGCTCTGAAATCAAAACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAACGATCCAGTGGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAGTGTATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACCAAGT 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTCAATGGGAAGTTCTTGGCGCATGCTTAATATCCAAAGTTGATGAAGTCAGAGAG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGCTATATTTTGAAGCCACCAAGCTCCCTTTTAGAGCATCCTCTAGTACTAGTCACT 1833
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TACGTAATCTCTGAGAGGTGACAAGCTGACTGACCGTGGCTACTCATCTTCTGCTGC 1893
DB 541 TyrValAsnProGluGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGACATGTGCTTTTATTAAGTATAGTAACACAGAGAAATCCACACTGT 1953
DB 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTCTTACAGCTTATCAAGTCTCTGAAGATGACCCCAACTTGGCAAAACAAAGGAATTT 2013
DB 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
```

```
QY 2014 TGGCCACCAATTTTGGATTGAGGCTCTCTCTGACTACTATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGACTTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAGAAATATCTACTGTGCTGTTTCAATATGTTGGTCTCTCAGGTGCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTTGTAGTATAGACACAGGGGATCTCTCACCAGGGCTTAAATTTGAGGC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTCATTTGACTTGTAGTCTGTGGGATCCAGGCTGCTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
QY 2434 GGAGGTACCTCTCCCTGATGGCAATTAATGCAGAGTGCAGATATCTTCAGGGTTCATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACAGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCAACCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATCCAGCAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGAAATGC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTCACATACCACTATATTACTAGTGTGTTTTAGTGGGCTGGAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGAGAGACACAGCATAAGAGTTTCTTGAATCGGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATGATCTTTTGCTACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 2
; Sequence 10, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: 08842.0019
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 16
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-10
Alignment Scores: 8.02e-16 Length: 24
Pred. No.: 24.00 Matches: 24
Score: 24.00
```

```
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 10
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-10
Alignment Scores: 4.87e-104 Length: 109
Pred. No.: 109.00 Matches: 109
Score: 109.00 Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 10.6% Gaps: 0
Db: 11

US-10-825-632-2 (1-3120) x US-11-176-951-10 (1-109)
QY 2533 GGATACACGGAACCTTATATGGGTCACTGACCTGACCAAGATGAACAGGGCTATTACTTAGGA 2592
Db 1 GlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGly 20
QY 2593 TCTGTGCCATGCAAGCAGCAAAAGTTCCCTCTCTGAAACCAAAATCGTTTACTGCTCTTACAT 2652
Db 21 SerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHis 40
QY 2653 GGTTCCTCGATGAGATGTCATTTTGCATATCCAGTATATCTGATTTTACTGATTTTACTGT 2712
Db 41 GlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuVal 60
QY 2713 AGGCTCGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTT 2772
Db 61 ArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgVal 80
QY 2773 CCTGAATCGGAGAACATTAATGAATGCACTCTTTTGCATCTTTTGCATCTTCAAGAAACCTTGA 2832
Db 81 ProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGly 100
QY 2833 TCAGTATTGCTGCTCTAAAAGTGATA 2859
Db 101 SerArgIleAlaAlaLeuLysValIle 109

RESULT 3
US-11-176-951-16
; Sequence 16, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: 08842.0019
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 16
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-16
Alignment Scores: 8.02e-16 Length: 24
Pred. No.: 24.00 Matches: 24
Score: 24.00
```

Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 2.3% Indels: 0  
 DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-176-951-16 (1-24)

QY 1099 GAAATATTATTCATCCCTATCTTGAACAGGAGGCGAGATTCATTCGGTTAT 1158  
 |||||  
 Db 1 GluIleIleHisValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyr 20  
 |||||

QY 1159 CCTAAACAGGT 1170  
 |||||  
 Db 21 ProLysThrGly 24

RESULT 4  
 US-11-176-951-11  
 ; Sequence 11, Application US/11176951  
 ; Publication No. US2006024313A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHEN, XIN  
 ; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF  
 ; FILE REFERENCE: 08842.0019  
 ; CURRENT APPLICATION NUMBER: US/11/176,951  
 ; PRIOR FILING DATE: 2005-07-06  
 ; PRIOR APPLICATION NUMBER: 60/586,095  
 ; PRIOR FILING DATE: 2004-07-06  
 ; PRIOR APPLICATION NUMBER: 60/585,952  
 ; PRIOR FILING DATE: 2004-07-06  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn Ver. 3.3  
 ; SEQ ID NO 11  
 ; LENGTH: 99  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-11-176-951-11

Alignment Scores:  
 Pred. No.: 0.0245 Length: 99  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 1.1% Indels: 0  
 DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-176-951-11 (1-99)

QY 2647 TTACATGTTTCCTGGATGAGATGCCATTTT 2679  
 |||||  
 Db 39 LeuHisGlyPheLeuAspGluAsnValHisPhe 49

RESULT 5  
 US-11-264-096-1591  
 ; Sequence 1591, Application US/11264096  
 ; Publication No. US20060084794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546D1  
 ; CURRENT APPLICATION NUMBER: US/11/264,096  
 ; PRIOR FILING DATE: 2005-11-02  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1

US-11-264-096-1591  
 ; Sequence 1593, Application US/11264096  
 ; Publication No. US20060084794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546D1  
 ; CURRENT APPLICATION NUMBER: US/11/264,096  
 ; PRIOR FILING DATE: 2005-11-02  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1

US-11-264-096-1593  
 ; Sequence 1593, Application US/11264096  
 ; Publication No. US20060084794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546D1  
 ; CURRENT APPLICATION NUMBER: US/11/264,096  
 ; PRIOR FILING DATE: 2005-11-02  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1

US-11-264-096-1593  
 ; Sequence 1593, Application US/11264096  
 ; Publication No. US20060084794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546D1  
 ; CURRENT APPLICATION NUMBER: US/11/264,096  
 ; PRIOR FILING DATE: 2005-11-02  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1

US-11-264-096-1593  
 ; Sequence 1593, Application US/11264096  
 ; Publication No. US20060084794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546D1  
 ; CURRENT APPLICATION NUMBER: US/11/264,096  
 ; PRIOR FILING DATE: 2005-11-02  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1

US-11-264-096-1593  
 ; Sequence 1593, Application US/11264096  
 ; Publication No. US20060084794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546D1  
 ; CURRENT APPLICATION NUMBER: US/11/264,096  
 ; PRIOR FILING DATE: 2005-11-02  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1

US-11-264-096-1593  
 ; Sequence 1593, Application US/11264096  
 ; Publication No. US20060084794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546D1  
 ; CURRENT APPLICATION NUMBER: US/11/264,096  
 ; PRIOR FILING DATE: 2005-11-02  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1

US-11-264-096-1593  
 ; Sequence 1593, Application US/11264096  
 ; Publication No. US20060084794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546D1  
 ; CURRENT APPLICATION NUMBER: US/11/264,096  
 ; PRIOR FILING DATE: 2005-11-02  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21  
 ; PRIOR APPLICATION NUMBER: 60/199,384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 2267  
 ; SOFTWARE: PatentIn Ver. 2.1

US-11-264-096-1593  
 ; Sequence 1593, Application US/11264096  
 ; Publication No. US20060084794A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF546D1  
 ; CURRENT APPLICATION NUMBER: US/11/264,096  
 ; PRIOR FILING DATE: 2005-11-02  
 ; PRIOR APPLICATION NUMBER: 09/833,245  
 ; PRIOR FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931

NAME/KEY: SITE  
LOCATION: (183)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-11-264-096-1593

Alignment Scores:  
Pred. No.: 0.0241 Length: 497  
Score: 11.00 Matches: 11  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 1.1% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-264-096-1593 (1-497)

QY 1804 CCTTAGAGCATCACCTGACGTAGTCAGTTAC 1836  
Db 156 ProlenGluHisLeuTyValValSerTy 166

RESULT 7

US-11-264-096-1594  
Sequence 1594, Application US/11264096  
Publication No. US20060084794A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PF546D1  
CURRENT APPLICATION NUMBER: US/11/264,096  
CURRENT FILING DATE: 2005-11-02  
PRIOR APPLICATION NUMBER: 09/833,245  
PRIOR FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: PatentIn ver. 2.1  
SEQ ID NO 1594  
LENGTH: 497  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-11-264-096-1594

Alignment Scores:  
Pred. No.: 0.0241 Length: 497  
Score: 11.00 Matches: 11  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 1.1% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-264-096-1594 (1-497)

QY 1804 CCTTAGAGCATCACCTGACGTAGTCAGTTAC 1836  
Db 156 ProlenGluHisLeuTyValValSerTy 166

RESULT 8

US-11-172-740-719  
Sequence 719, Application US/11172740  
Publication No. US2006005724A1  
GENERAL INFORMATION:  
APPLICANT: MASCIA, Peter  
APPLICANT: ALEXANDROV, Nikolai  
APPLICANT: BROVER, Vyacheslav  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FOR  
TITLE OF INVENTION: PLANT CHARACTERISTICS AND PHENOTYPES  
FILE REFERENCE: 2750-1602PUS2  
CURRENT APPLICATION NUMBER: US/11/172,740  
CURRENT FILING DATE: 2005-06-30  
PRIOR APPLICATION NUMBER: 60/593,621

PRIOR FILING DATE: 2004-06-30  
PRIOR APPLICATION NUMBER: 60/584,829  
PRIOR FILING DATE: 2004-06-30  
PRIOR APPLICATION NUMBER: 60/584,800  
PRIOR FILING DATE: 2004-06-30  
NUMBER OF SEQ ID NOS: 2523  
SEQ ID NO 719  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)-(353)  
OTHER INFORMATION: Public GI no. 50906397  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION:  
OTHER INFORMATION: Utility: Useful for delaying flowering time  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION:  
OTHER INFORMATION: Utility: Useful for increasing chlorophyll and photosynthetic cap.

QY 23 AAGCGCTGCTACTGCGCGCTGCT 49  
Db 313 LysAlaAlaThraAlaAlaAla 321

RESULT 9

US-11-071-259-3  
Sequence 3, Application US/11071259  
Publication No. US2006008440A1  
GENERAL INFORMATION:  
APPLICANT: Blatt, Thomas  
APPLICANT: Mummert, Christopher  
APPLICANT: Mundt, Claudia  
APPLICANT: Staeb, Franz  
TITLE OF INVENTION: Cosmetic or Dermatological Preparations Having a Content of  
FILE REFERENCE: P26688  
CURRENT APPLICATION NUMBER: US/11/071,259  
CURRENT FILING DATE: 2005-03-04  
PRIOR APPLICATION NUMBER: PCT/EP2002/010044  
PRIOR FILING DATE: 2002-09-07  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 3  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Limanda ferruginea  
US-11-071-259-3

Alignment Scores:  
Pred. No.: 2.88 Length: 353  
Score: 9.00 Matches: 9  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.9% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-172-740-719 (1-353)

QY 23 AAGCGCTGCTACTGCGCGCTGCT 49  
Db 313 LysAlaAlaThraAlaAlaAla 321

Alignment Scores:  
Pred. No.: 32 Length: 49  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-071-259-3 (1-49)

Qy 26 GCTGCTACTGCGCGCTGCT 49  
Db 10 AlaAlaAlaThrAlaAlaAla 17

## RESULT 10

US-11-096-568A-9319  
; Sequence 9319, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 9319  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(90)  
; OTHER INFORMATION: Ceres Seq. ID no. 15226318  
US-11-096-568A-9319

Alignment Scores:  
Pred. No.: 31.8 Length: 90  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-096-568A-9319 (1-90)

Qy 48 GCAGCGCGCGCTAGCAGCGCC 25  
Db 45 AlaAlaAlaValaAlaAla 52

## RESULT 11

US-10-504-588-6  
; Sequence 6, Application US/10504588  
; Publication No. US20060078498A1  
; GENERAL INFORMATION:  
; APPLICANT: Buckholz, Thomas  
; APPLICANT: Vandenberg, Mark  
; APPLICANT: Heitmeier, Stefan  
; APPLICANT: Taylor, Ian  
; APPLICANT: Gedrich, Richard  
; TITLE OF INVENTION: Methods for the Identification of Novel Ligands for the G  
; FILE REFERENCE: 5121-PCT  
; CURRENT APPLICATION NUMBER: US/10/504,588  
; CURRENT FILING DATE: 2004-08-12  
; PRIOR APPLICATION NUMBER: PCT/US03/09522  
; PRIOR FILING DATE: 2003-04-28  
; PRIOR APPLICATION NUMBER: US 60/368,849  
; PRIOR FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 6

; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-504-588-6  
Alignment Scores:  
Pred. No.: 31.8 Length: 108  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 9 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-504-588-6 (1-108)

Qy 30 CTGCTACTGCGCGCTGCTTCTT 53  
Db 13 LeuLeuLeuProLeuLeuLeu 20

## RESULT 12

US-11-073-420-6  
; Sequence 6, Application US/11073420  
; Publication No. US20060019338A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhou, Qun-Yong  
; TITLE OF INVENTION: Primate Prokineticin and Prokineticin  
; FILE REFERENCE: UC11210-1  
; CURRENT APPLICATION NUMBER: US/11/073,420  
; CURRENT FILING DATE: 2005-03-04  
; PRIOR APPLICATION NUMBER: 60/550,753  
; PRIOR FILING DATE: 2004-03-05  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Macaca mulatta  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)..(27)  
US-11-073-420-6

Alignment Scores:  
Pred. No.: 31.8 Length: 108  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-073-420-6 (1-108)

Qy 30 CTGCTACTGCGCGCTGCTTCTT 53  
Db 13 LeuLeuLeuProLeuLeuLeu 20

## RESULT 13

US-11-194-246-399  
; Sequence 399, Application US/11194246  
; Publication No. US20050272069A1  
; GENERAL INFORMATION:  
; APPLICANT: Mott, John  
; APPLICANT: Trepod, Catherine  
; APPLICANT: Arvidson, Staifan  
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND METI  
; FILE REFERENCE: 00592.US1 (M&R 368.05920101)  
; CURRENT APPLICATION NUMBER: US/11/194,246  
; CURRENT FILING DATE: 2005-08-01  
; PRIOR APPLICATION NUMBER: US/10/274,586  
; PRIOR FILING DATE: 2002-10-21

;; PRIOR APPLICATION NUMBER: US 60/345,438  
;; PRIOR FILING DATE: 2001-10-19  
;; NUMBER OF SEQ ID NOS: 621  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 399  
;; LENGTH: 123  
;; TYPE: PRT  
;; ORGANISM: HAEMOPHILUS INFLUENZAE  
US-11-194-246-399

Alignment Scores:  
Pred. No.: 31.7 Length: 123  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-194-246-399 (1-123)

Qy 48 GCAGCGCGCGCAGTAGCAGCGGCC 25  
Db 35 AlaAlaAlaAlaValAlaAlaAla 42

RESULT 14

US-11-096-568A-14806  
; Sequence 14806, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 14806  
; LENGTH: 129  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mays  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(129)  
; OTHER INFORMATION: Ceres Seq. ID no. 11415220  
US-11-096-568A-14806

Alignment Scores:  
Pred. No.: 31.7 Length: 129  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-096-568A-14806 (1-129)

Qy 48 GCAGCGCGCGCAGTAGCAGCGGCC 25  
Db 5 AlaAlaAlaAlaValAlaAlaAla 12

RESULT 15

US-11-096-568A-26015  
; Sequence 26015, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 26015

;; LENGTH: 129  
;; TYPE: PRT  
;; ORGANISM: Zea mays subsp. mays  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1)..(129)  
;; OTHER INFORMATION: Ceres Seq. ID no. 13498137  
US-11-096-568A-26015

Alignment Scores:  
Pred. No.: 31.7 Length: 129  
Score: 8.00 Matches: 8  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 0.8% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-096-568A-26015 (1-129)

Qy 48 GCAGCGCGCGCAGTAGCAGCGGCC 25  
Db 5 AlaAlaAlaAlaValAlaAlaAla 12

Search completed: May 2, 2006, 05:01:48  
Job time : 69.5 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:41:02 ; Search time 1064.84 Seconds  
(without alignments)  
9295.694 Million cell updates/sec

Title: US-10-825-632-4

Perfect score: 1197

Sequence: 1 attttgaggcaccacaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA\_Main:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1197	100.0	1197	8	US-10-825-632-4
2	1179	98.5	4676	3	US-09-976-674-20
3	1179	98.5	4676	3	US-10-982-512-20
4	1040	86.9	3143	6	US-10-170-789-37
5	1036	86.5	3120	7	US-10-415-122-5
6	1036	86.5	3120	8	US-10-825-632-2
7	1033.2	86.3	4309	3	US-09-976-674-14
8	1033.2	86.3	4309	9	US-10-982-512-14
9	1033.2	86.3	4829	3	US-09-976-674-12
10	1033.2	86.3	4829	9	US-10-982-512-12
11	1033	86.3	3106	7	US-10-311-035-30
12	1013.6	84.7	4685	3	US-09-976-674-22
13	1013.6	84.7	4685	9	US-10-982-512-22
14	881	73.6	1669	8	US-10-825-632-6
15	881	73.6	2830	3	US-10-956-157-2177
16	873	72.9	4523	3	US-09-976-674-8
17	873	72.9	4523	9	US-10-982-512-8
18	863.6	72.1	2510	7	US-10-275-505-16
19	863.6	72.1	2510	10	US-11-140-224-16
20	791.4	66.1	925	6	US-10-264-237-710
21	778	65.0	2649	6	US-10-170-789-39
22	776.4	64.9	2649	6	US-10-054-776-1
23	776.4	64.9	2671	3	US-09-976-674-2

24 776.4 64.9 2671 9 US-10-982-512-2 Sequence 2, Appli  
25 594 49.6 600 9 US-10-956\*157-7412 Sequence 7412, Ap  
26 459 38.3 1083 8 US-10-825-632-8 --- Sequence 8, Appli  
27 380.4 31.8 561 3 US-09-764-891-877 Sequence 877, App  
28 308.8 25.8 502 3 US-09-918-995-19585 Sequence 19585, A  
29 237.2 19.8 3287 7 US-10-415-122-3 Sequence 3, Appli  
30 214.4 17.9 2660 7 US-10-072-012-223 Sequence 223, App  
31 214.4 17.9 2660 7 US-10-072-012-225 Sequence 225, App  
32 212.8 17.8 2495 7 US-10-415-122-8 Sequence 8, Appli  
33 212.8 17.8 2617 3 US-09-976-674-4 Sequence 4, Appli  
34 212.8 17.8 2617 9 US-10-982-512-4 Sequence 4, Appli  
35 212.8 17.8 3000 7 US-10-415-122-1 Sequence 1, Appli  
36 212.8 17.8 3716 9 US-10-433-757-30 Sequence 30, Appli  
37 212.8 17.8 4180 3 US-09-976-674-36 Sequence 36, Appli  
38 212.8 17.8 4180 9 US-10-982-512-36 Sequence 36, Appli  
39 212.8 17.8 4219 3 US-09-976-674-28 Sequence 28, Appli  
40 212.8 17.8 4219 9 US-10-982-512-28 Sequence 28, Appli  
41 212.8 17.8 4263 3 US-09-976-674-34 Sequence 34, Appli  
42 212.8 17.8 4263 9 US-10-982-512-34 Sequence 34, Appli  
43 212.8 17.8 4302 3 US-09-976-674-24 Sequence 24, Appli  
44 212.8 17.8 4302 9 US-10-982-512-24 Sequence 24, Appli  
45 211.2 17.6 2261 6 US-10-094-749-47 Sequence 47, Appli

#### ALIGNMENTS

RESULT 1

US-10-825-632-4  
; Sequence 4, Application US/10825632  
; Publication No. US20040191826A1  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: FCSB-100-Div. 1  
; CURRENT APPLICATION NUMBER: US/10/825,632  
; CURRENT FILING DATE: 2004-04-15  
; PRIOR APPLICATION NUMBER: US 10/070,464  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 1197  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; US-10-825-632-4

Query Match 100.0%; Score 1197; DB 8; Length 1197;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1197; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 ATTTTGAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTACGTAA 60  
Db 1 ATTTTGAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTAA 60  
Qy 61 ATCTGGAGAGGTGACAAAGGTGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 120  
Db 61 ATCTGGAGAGGTGACAAAGGTGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 120  
Qy 121 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACATGTGTGTCCC 180  
Db 121 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACATGTGTGTCCC 180  
Qy 181 TTTCAGCTATCAAGTCCCTGAAGATGACCCAACTTTGGGAAAAAAAGGAATTTTGGGCA 240  
Db 181 TTTCAGCTATCAAGTCCCTGAAGATGACCCAACTTTGGGAAAAAAAGGAATTTTGGGCA 240

QY 241 CCATTTGGATTACAGAGGCTCTCTCCCTGACTATACCTCCAGAAATTTCTCTTTG 300  
DB 241 CCAATTTGGATTACAGAGGCTCTCTCTCCCTGACTATACCTCCAGAAATTTCTCTTTG 300  
QY 301 AAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAAGGCTCATGATCTACAGGCTG 360  
DB 301 AAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAAGGCTCATGATCTACAGGCTG 360  
QY 361 GAAGAATAATCTCTACTGTGCTGTCTATATATGGTGGTCTCTCAGGGTCAAAATAGAAATG 420  
DB 361 GAAGAATAATCTCTACTGTGCTGTCTATATATGGTGGTCTCTCAGGGTCAAAATAGAAATG 420  
QY 421 ACGATCAGGTGGAAGGACTCAATATATCTAGCTCTCTCGATATGATTTTCACTTAGATC 480  
DB 421 ACGATCAGGTGGAAGGACTCAATATATCTAGCTCTCTCGATATGATTTTCACTTAGATC 480  
QY 481 GTGTGGGATCCAGGCTGTCTCTATGGGAGTACCTCTCCCTGTATGGCATTAATGAGA 540  
DB 481 GTGTGGGATCCAGGCTGTCTCTATGGGAGTACCTCTCCCTGTATGGCATTAATGAGA 540  
QY 541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATG 600  
DB 541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATG 600  
QY 601 ATACAGGATACACGGACGTTTATATGGGTACCCCTGACAGAAATGAACAGGCTATTACT 660  
DB 601 ATACAGGATACACGGACGTTTATATGGGTACCCCTGACAGAAATGAACAGGCTATTACT 660  
QY 661 TAGGATCTGTGGCCATCGACAGCAAGGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720  
DB 661 TAGGATCTGTGGCCATCGACAGCAAGGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720  
QY 721 TACATGGTTTCTCGATGAGAAATGTCGATTTTGACATACCAAGTATATTACTGATTTT 780  
DB 721 TACATGGTTTCTCGATGAGAAATGTCGATTTTGACATACCAAGTATATTACTGATTTT 780  
QY 781 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAA 840  
DB 781 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAA 840  
QY 841 GAGTTCTGAAATCGGGAGAACATTTATGAATGATCTCTTTTGCACTACCTTCAAGAAACC 900  
DB 841 GAGTTCTGAAATCGGGAGAACATTTATGAATGATCTCTTTTGCACTACCTTCAAGAAACC 900  
QY 901 TTGGATCACGTTATGCTGCTCTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGG 960  
DB 901 TTGGATCACGTTATGCTGCTCTAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGG 960  
QY 961 TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATC 1020  
DB 961 TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATC 1020  
QY 1021 ATACATTTTGTATACCTGCCATGTAACATCTACTCTGAAATAAATGTTGGTGCATGCA 1080  
DB 1021 ATACATTTTGTATACCTGCCATGTAACATCTACTCTGAAATAAATGTTGGTGCATGCA 1080  
QY 1081 GGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAACCCACATGCTCAAAATCAAAATG 1140  
DB 1081 GGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAACCCACATGCTCAAAATCAAAATG 1140  
QY 1141 ATACATATCTCTGAGAGACCCAGCAATACCAATAGAAATTAATCAAAAAA 1197  
DB 1141 ATACATATCTCTGAGAGACCCAGCAATACCAATAGAAATTAATCAAAAAA 1197

RESULT 2  
US-09-976-674-20  
; Sequence 20, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-976-674-20  
  
Query Match 98.5%; Score 1179; DB 3; Length 4676;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1193; Conservative 0; Mismatches 0; Indels 4; Gaps 1;  
  
QY 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAA 60  
DB 1781 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 1840  
QY 61 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCATTCTTGTGTCATGTC 120  
DB 1841 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCATTCTTGTGTCATGTC 1900  
QY 121 AGCACTGTGACTTCTTTATAGTAAGTATAGTAAGTAAACAGAAATCCACACTGTGTGTC 180  
DB 1901 AGCACTGTGACTTCTTTATAGTAAGTATAGTAAGTAAACAGAAATCCACACTGTGTGTC 1960  
QY 181 TTTTCAAGCTATCAAGTCTGGAAGTACCCAACTTGGAAACAAAGAAATTTTGGGCA 240  
DB 1961 TTTTCAAGCTATCAAGTCTGGAAGTACCCAACTTGGAAACAAAGAAATTTTGGGCA 2020  
QY 241 CCATTTTGGATTCAGAGGTCCTCTCTGACTATCTCTCCAGAAATTTTCTCTTTG 300  
DB 2021 CCATTTTGGATTCAGAGGTCCTCTCTGACTATCTCTCCAGAAATTTTCTCTTTG 2080  
QY 301 AAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAAGGCTCATGATCTACAGCCTG 360  
DB 2081 AAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAAGGCTCATGATCTACAGCCTG 2140  
QY 361 GAAAGAAATATCCTTACTGCTCTTATATATATGGTGGTCTCAGGGTCAAAATAGAAATG 420  
DB 2141 GAAAGAAATATCCTTACTGCTCTTATATATATGGTGGTCTCAGGGTCAAAATAGAAATG 2196  
QY 421 ACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCACTTAGATC 480  
DB 2197 ACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCACTTAGATC 2256  
QY 481 GTGTGGGATCCAGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGAGA 540  
DB 2257 GTGTGGGATCCAGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGAGA 2316  
QY 541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATG 600  
DB 2317 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATG 2376  
QY 601 ATACAGGATACAGGACGTTTATATGGGTCAACCTGACAGAAATGAACAGGCTATTACT 660  
DB 2377 ATACAGGATACAGGACGTTTATATGGGTCAACCTGACAGAAATGAACAGGCTATTACT 2436  
QY 661 TAGGATCTGTGGCCATGCAAGGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720  
DB 2437 TAGGATCTGTGGCCATGCAAGGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 2496  
QY 721 TACATGGTTTCTCGATGAGAAATGTCATTTTGGCATACACAGTATATATTCTGAGTTTTT 780  
DB 2497 TACATGGTTTCTCGATGAGAAATGTCATTTTGGCATACACAGTATATATTCTGAGTTTTT 2556  
QY 781 TAGTGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATAA 840



```

1 APPLICANT: Curtis , Rory A. J.
2 APPLICANT: Williamson, Mark
3 APPLICANT: Welch, Nadine
4 TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
5 TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
6 FILE REFERENCE: 10448-191001
7 CURRENT APPLICATION NUMBER: US/10/170,789
8 CURRENT FILING DATE: 2002-06-13
9 PRIOR APPLICATION NUMBER: US 09/797,039
10 PRIOR FILING DATE: 2001-02-28
11 PRIOR APPLICATION NUMBER: PCT/US01/06525
12 PRIOR FILING DATE: 2001-02-28
13 PRIOR APPLICATION NUMBER: US 60/186,061
14 PRIOR FILING DATE: 2000-02-29
15 PRIOR APPLICATION NUMBER: US 09/882,166
16 PRIOR FILING DATE: 2001-06-15
17 PRIOR APPLICATION NUMBER: PCT/US01/19269
18 PRIOR FILING DATE: 2001-06-15
19 PRIOR APPLICATION NUMBER: US 60/212,078
20 PRIOR FILING DATE: 2000-06-15
21 PRIOR APPLICATION NUMBER: US 09/934,406
22 PRIOR FILING DATE: 2001-08-21
23 PRIOR APPLICATION NUMBER: PCT/US01/26052
24 PRIOR FILING DATE: 2001-08-21
25 PRIOR APPLICATION NUMBER: US 60/226,740
26 PRIOR FILING DATE: 2000-08-21
27 PRIOR APPLICATION NUMBER: US 09/861,801
28 PRIOR FILING DATE: 2001-05-21
29 PRIOR APPLICATION NUMBER: PCT/US01/16549
30 PRIOR FILING DATE: 2001-05-21
31 PRIOR APPLICATION NUMBER: US 60/205,508
32 PRIOR FILING DATE: 2000-05-19
33 PRIOR APPLICATION NUMBER: US 09/801,267
34 PRIOR FILING DATE: 2001-03-06
35 PRIOR APPLICATION NUMBER: PCT/US01/07138
36 PRIOR FILING DATE: 2001-03-05
37 PRIOR APPLICATION NUMBER: US 60/187,454
38 PRIOR FILING DATE: 2000-03-07
39 PRIOR APPLICATION NUMBER: US 09/829,671
40 PRIOR FILING DATE: 2001-04-10
41 PRIOR APPLICATION NUMBER: PCT/US01/40483
42 PRIOR FILING DATE: 2001-04-11
43 PRIOR APPLICATION NUMBER: US 60/197,508
44 PRIOR FILING DATE: 2000-04-18
45 PRIOR APPLICATION NUMBER: US 09/961,721
46 PRIOR FILING DATE: 2001-09-24
47 PRIOR APPLICATION NUMBER: PCT/US01/29904
48 PRIOR FILING DATE: 2001-09-24
49 PRIOR APPLICATION NUMBER: US 60/235,023
50 PRIOR FILING DATE: 2000-09-25
51 PRIOR APPLICATION NUMBER: US 10/045,367
52 PRIOR FILING DATE: 2001-11-07
53 PRIOR APPLICATION NUMBER: US 60/246,561
54 PRIOR FILING DATE: 2000-11-07
55 PRIOR APPLICATION NUMBER: US 09/801,275
56 PRIOR FILING DATE: 2001-03-06
57 PRIOR APPLICATION NUMBER: PCT/US01/07074
58 PRIOR FILING DATE: 2001-03-05
59 PRIOR APPLICATION NUMBER: US 60/187,420
60 PRIOR FILING DATE: 2000-03-07
61 NUMBER OF SEQ ID NOS: 63
62 SOFTWARE: FastSeq for Windows version 4.0
63 SEQ ID NO 37
64 LENGTH: 3143
65 TYPE: DNA
66 ORGANISM: Homo sapiens
67 FEATURE:
68 NAME/KEY: CDS
69 LOCATION: (229)...(2874)
70 US-10-170-789-37

```

Query Match	86.9%	Score 1040;	DB 6;	Length 3143;
Best Local Similarity	89.1%	Pred. No. 8.6e-303;		

Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;									
Qy	1	ATTTTGAGGACCAAGAAGCTCCCTTTAGAGCATACCTGTACGTAGTACGTTACGTAA	60						
Db	1796	ATTTTGAGGACCAAGAAGCTCCCTTTAGAGCATACCTGTACGTAGTACGTTACGTAA	1855						
Qy	61	ATCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGCATCAGTC	120						
Db	1856	ATCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGCATCAGTC	1915						
Qy	121	AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTC	180						
Db	1916	AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACACAGAAATCCACACTGTGTGTC	1975						
Qy	181	TTTACAAGCTATCAAGTCTCAAGATGACCAACTTGC AAAACAAGGAATTTTGGGCA	240						
Db	1976	TTTACAAGCTATCAAGTCTCAAGATGACCAACTTGC AAAACAAGGAATTTTGGGCA	2035						
Qy	241	CCATTTTGGATTACAGCGTCTCTCTGACATATACTCTCAGAAATTTTCTTTTG	300						
Db	2036	CCATTTTGGATTACAGCGTCTCTCTGACATATACTCTCAGAAATTTTCTTTTG	2095						
Qy	301	AAAGTACTACTGGATTACATGTATGGATGCTCTACAAGCCTCATGATCTACAGCTG	360						
Db	2096	AAAGTACTACTGGATTACATGTATGGATGCTCTACAAGCCTCATGATCTACAGCTG	2155						
Qy	361	GAAGGAATATCCTACTGTGCTGTTCTATATATGGTGGTCTCA-----	403						
Db	2156	GAAGGAATATCCTACTGTGCTGTTCTATATATGGTGGTCTCAAGTGCAGTTGGTGAATA	2215						
Qy	404	-----	403						
Db	2216	ATCGGTTTAAAGGATCAAGTATTTCCGCTTGAATACCTTAGCCTCTTAGGTTATGTGG	2275						
Qy	404	-----	403						
Db	2276	TTGTAGTGATAGACAACAGGGGATCCTGTCAACGAGGCTTAAATTTGAAGCGCCTTTA	2335						
Qy	404	-----GGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT	453						
Db	2336	AAATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT	2395						
Qy	454	CTCGATATGATTTTCATTGACTTATAGATCTGTGTGGGATCCACGGCTGGTCTTATGGAGGAT	513						
Db	2396	CTCGATATGATTTTCATTGACTTATAGATCTGTGTGGGATCCACGGCTGGTCTTATGGAGGAT	2455						
Qy	514	ACCTCTCCCTCATGGCAATTAATCGACAGTCAAGATATCTTCAGAGGTTGCTATTGCTGGGG	573						
Db	2456	ACCTCTCCCTCATGGCAATTAATCGACAGTCAAGATATCTTCAGAGGTTGCTATTGCTGGGG	2515						
Qy	574	CCCAAGTCACATCTGTGATCTCTATGATACAGGATACACGACAGTTATTCGGTCACC	633						
Db	2516	CCCAAGTCACATCTGTGATCTCTATGATACAGGATACACGACAGTTATTCGGTCACC	2575						
Qy	634	CTGACCAAGATGACAGGCTATTACTTAGGATCTGTGGCCATCCGACAGCAAAAAGTTCC	693						
Db	2576	CTGACCAAGATGACAGGCTATTACTTAGGATCTGTGGCCATCCGACAGCAAAAAGTTCC	2635						
Qy	694	CCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTTG	753						
Db	2636	CCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTTG	2695						
Qy	754	CACATACAGTATATTACTAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGA	813						
Db	2696	CACATACAGTATATTACTAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGA	2755						
Qy	814	TCTATCTCAGGAGAGACACACATAAAGGTTCTCGAATCGGAGAACATATATGAATGC	873						
Db	2756	TCTATCTCAGGAGAGACACACATAAAGGTTCTCGAATCGGAGAACATATATGAATGC	2815						
Qy	874	ATCTTTTGCACTACCTTCAAGAAAACTTTGGATCACGTTATGTGCTTAAAAGTGATAT	933						
Db	2816	ATCTTTTGCACTACCTTCAAGAAAACTTTGGATCACGTTATGTGCTTAAAAGTGATAT	2875						

934	QY		AATTTTGACCTGCTGTGAACATCTCTCTGGTATACATCTGGCTATTATTAACCAAATGAGGAGGTT	993
2876	Db		AATTTTGACCTGCTGTGAGAACTCTCTGGTATACATCTGGCTATTATTAACCAAATGAGGAGGTT	2935
994	QY		TAACTCAACAGAAAAACACAGAAATTTGATCATCATATTTTGATACCTTGCATCTGCAATGTAACATCTAC	1053
2936	Db		TAACTCAACAGAAAAACACAGAAATTTGATCATCATATTTTGATACCTTGCATCTGCAATGTAACATCTAC	2995
1054	QY		TCTTGAAATTAATGTGTGGTGGCCATGACAGGGGCTACGGTTTGTGGTAGTAAATCTAATATACC	1113
2996	Db		TCTTGAAATTAATGTGTGGTGGCCATGACAGGGGCTACGGTTTGTGGTAGTAAATCTAATATACC	3055
1114	QY		TTAATCCCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATATACCATA	1173
3056	Db		TTAATCCCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATATACCATA	3115
1174	QY		AGAATTACTAAAAAATAAAAAATAAAAA	1197
3116	Db		AGAATTACTAAAAAATAAAAAATAAAAA	3139

## RESULT 5

```

US-10-415-122-5
; Sequence 5, Application US/10415122
; Publication NO. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FPI5217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-415-122-5

```

Query Match : 86.5%; Score 1036; DB 7; Length 3120;  
Best Local Similarity 89.0%; Pred. No. 1.4e-301;  
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

Qy	1	ATTTTGAAGGCACCAAGACTCCCTTTTAGAGATCACTGTACTGTACGTACGTAA	60
Db	1781	ATTTTGAAGGCACCAAGACTCCCTTTTAGAGATCACTGTACTGTACGTACGTAA	1840
Qy	61	ATCTCGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAATTCTTGTGTCATCATGTC	120
Db	1841	ATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAATTCTTGTGTCATCATGTC	1900
Qy	121	AGCACTGTGACTTCTTTTAAGTAAGTATAGTAAACAGAGAAGATCCACACTGTGTGTCCC	180
Db	1901	AGCACTGTGACTTCTTTTAAGTAAGTATAGTAAACAGAGAAGATCCACACTGTGTGTCCC	1960
Qy	181	TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTTGCAAAACAAGGAATTTTGGGCCA	240
Db	1961	TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTTGCAAAACAAGGAATTTTGGGCCA	2020
Qy	241	CAATTTTGGATTACGACAGTCCCTCTCTGACTATACTCTCTCAGAAATTTTCTCTTTTG	300
Db	2021	CAATTTTGGATTACGACAGTCCCTCTCTGACTATACTCTCTCAGAAATTTTCTCTTTTG	2080
Qy	301	AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGGCTCATGATCTACAGCCTG	360
Db	2081	AAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGGCTCATGATCTACAGCCTG	2140
Qy	361	GAAGAAATATCCTACTGTGCTGTTTCATATATATGGTGGTCTCTCA-----	403
Db	2141	GAAGAAATATCCTACTGTGCTGTTTCATATATGGTGGTCTCTCAGGTGCAAGTTGGTGAATA	2200
Qy	404	-----	403

```
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-2

Query Match      86.5%; Score 1036; DB 8; Length 3120;
Best Local Similarity 89.0%; Pred. No. 1.4e-301;
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAGTTAGGTAA 60
DB 1781 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAGTTAGGTAA 1840

QY 61 ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTC 120
DB 1841 ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTC 1900

QY 121 AGCACTGTGACTTCTTTTAACTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180
DB 1901 AGCACTGTGACTTCTTTTAACTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 1960

QY 181 TTTCACAGCTATCAAGTCTGGAAGATGACCACTTGCACCAACCAAGAAATTTGGGCA 240
DB 1961 TTTCACAGCTATCAAGTCTGGAAGATGACCACTTGCACCAACCAAGAAATTTGGGCA 2020

QY 241 CCATTTTGGATTACAGAGTCTCTTCCTGATCATCTCTCCAGAAATTTCTCTTTG 300
DB 2021 CCATTTTGGATTACAGAGTCTCTTCCTGATCATCTCTCCAGAAATTTCTCTTTG 2080

QY 301 AAGTACTACTGGATTACATTTGATGGATGCTCTCAAGCCCTCATGATCTACAGCCGTG 360
DB 2081 AAGTACTACTGGATTACATTTGATGGATGCTCTCAAGCCCTCATGATCTACAGCCGTG 2140

QY 361 GAAAGAAATATCTTACTGCTGCTTCATATATGTTGCTCTCAAGGTCAGTTGGTGAATA 2200
DB 2141 GAAAGAAATATCTTACTGCTGCTTCATATATGTTGCTCTCAAGGTCAGTTGGTGAATA 2200

QY 404 ----- 403
DB 2201 ATCGGTTTAAAGAGTCAAGTATTTCCGCTTTGAATACCCCTAGCCTCTCTAGGTTATGTGG 2260

QY 404 ----- 403
DB 2261 TTGTAGTGATAGACAACAGGGGATCTCTGTCA CCGAGGGCTTAAATTTGAAGGCGCTTTTA 2320

QY 404 -----GGGTCAATAGAAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTT 453
DB 2321 AATAATAAATGGGTCAATAGAAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTT 2380

QY 454 CTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGAT 513
DB 2381 CTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGAT 2440

QY 514 ACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 573
DB 2441 ACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 2500

QY 574 CCCCACTGCTCTGTGGATCTTCTATGATACAGGATACCGAACTTATATGGGTGCTACC 633
DB 2501 CCCCACTGCTCTGTGGATCTTCTATGATACAGGATACCGAACTTATATGGGTGCTACC 2560
```

```
634 CTGACCAAGATGAACAGGGGCTATTACTTAGGATCTGTGGCCCATGCAACAGCAAAAGTTCC 693
2561 CTGACCAAGATGAACAGGGGCTATTACTTAGGATCTGTGGCCCATGCAACAGCAAAAGTTCC 2620
694 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTCTCTGGATGAGAATGTCCATTTTG 753
2621 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTCTCTGGATGAGAATGTCCATTTTG 2680
754 CACATACCAAGTATATTACTGAGTTTTTTAGTCAGGGCTGGAAAGCCATATGATTTACAGA 813
2681 CACATACCAAGTATATTACTGAGTTTTTTAGTCAGGGCTGGAAAGCCATATGATTTACAGA 2740
814 TCTATCTCTCAGGAGACACAGATAAGAGTTCTTGAATCGGGAGAACATTATCAACTGC 873
2741 TCTATCTCTCAGGAGACACAGATAAGAGTTCTTGAATCGGGAGAACATTATCAACTGC 2800
874 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTTATTGCTGCTCTAAAAGTGATAT 933
2801 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTTATTGCTGCTCTAAAAGTGATAT 2860
934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTT 993
2861 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTT 2920
994 TAATCAACAGAAAACACAGAAATTGATCATCATTTTGTATACCTGCCATTTAAACCAATGAGGAGTT 1053
2921 TAATCAACAGAAAACACAGAAATTGATCATCATTTTGTATACCTGCCATTTAAACCAATGAGGAGTT 2980
1054 TCCTGAAAATAAATGTGGTCCATCGAGGGCTCTACGGTTTGTGGTAGTAATCTTAATACC 1113
2981 TCCTGAAAATAAATGTGGTCCATCGAGGGCTCTACGGTTTGTGGTAGTAATCTTAATACC 3040
1114 TTAACCCACATCTCTCAAAATCAATGATCATATTTCTTGAGAGACCAGCAATACCATA 1173
3041 TTAACCCACATCTCTCAAAATCAATGATCATATTTCTTGAGAGACCAGCAATACCATA 3100
1174 AGAATTACTAAAAAAA 1193
3101 AGAATTACTAAAAAAA 3120

RESULT 7
US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 4309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-14

Query Match      86.3%; Score 1033.2; DB 3; Length 4309;
Best Local Similarity 88.7%; Pred. No. 1.2e-300;
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAGTTAGGTAA 60
DB 1261 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAGTTAGGTAA 1320
```





QY 403 ----- 402  
Db 1681 TAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTATGT 1740  
QY 403 ----- 402  
Db 1741 GGTGTAGTATAGACAACAGGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCCTT 1800  
QY 403 -----AGGTCAAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGC 451  
Db 1801 TAAATATAAATGGGTCAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGC 1860  
QY 452 TTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGG 511  
Db 1861 TTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGG 1920  
QY 512 ATACCTCTCCCTGATGSCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATGCTGG 571  
Db 1921 ATACCTCTCCCTGATGSCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATGCTGG 1980  
QY 572 GGCCCCAGTCACCTGTGGATCTTCTATGATACAGATACACGGAAAGCTTATATGGGTCA 631  
Db 1981 GGCCCCAGTCACCTGTGGATCTTCTATGATACAGATACACGGAAAGCTTATATGGGTCA 2040  
QY 632 CCCTGACCAAGTGAACGCGCTATTACTTAGATCTGTGGCATGCAAGCAGCAAAAGTT 691  
Db 2041 CCCTGACCAAGTGAACGCGCTATTACTTAGATCTGTGGCATGCAAGCAGCAAAAGTT 2100  
QY 692 CCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTTCTGATGAGATGTCATTT 751  
Db 2101 CCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTTCTGATGAGATGTCATTT 2160  
QY 752 TGCAATACAGATATATCTAGTGTGTTTTAGTAGGGCTGGAAAGCCATATGATTACA 811  
Db 2161 TGCAATACAGATATATCTAGTGTGTTTTAGTAGGGCTGGAAAGCCATATGATTACA 2220  
QY 812 GATCTATCTCAGAGAGACACAGCATAAAGATTCTGAAATCGGAGAACATTATGAAT 871  
Db 2221 GATCTATCTCAGAGAGACACAGCATAAAGATTCTGAAATCGGAGAACATTATGAAT 2280  
QY 872 GCATCTTTTGCATCTTCAAGAAACCTTGGATCAGTATTCGCTCTTAAAGTGAT 931  
Db 2281 GCATCTTTTGCATCTTCAAGAAACCTTGGATCAGTATTCGCTCTTAAAGTGAT 2340  
QY 932 ATAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAG 991  
Db 2341 ATAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAG 2400  
QY 992 TTTAATCAACAGAAAAACAGAAATGATCATCAATTTTGTATACCTGCCATGTAAACATCT 1051  
Db 2401 TTTAATCAACAGAAAAACAGAAATGATCATCAATTTTGTATACCTGCCATGTAAACATCT 2460  
QY 1052 ACTCCTGAAATATAATGTGGTGCCATGCAAGGGTCTACGGTTTGTGTAGTAAATCTAATA 1111  
Db 2461 ACTCCTGAAATATAATGTGGTGCCATGCAAGGGTCTACGGTTTGTGTAGTAAATCTAATA 2520  
QY 1112 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 1171  
Db 2521 CCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 2580  
QY 1172 TAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1197  
Db 2581 TAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2606

RESULT 9  
US-09-976-674-12  
; Sequence 12, Application US/09976674  
; Patent No. US20020115943A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 4829  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-976-674-12

Query Match 86.3%; Score 1033.2; DB 3; Length 4829;  
Best Local Similarity 88.7%; Pred. No. 1-2e-300;  
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGACTCCCTTTAGACATCACCTGTACGTAGTACGTACGTAA 60  
Db 1781 ATTTTGAAGGCACCAAGACTCCCTTTAGACATCACCTGTACGTAGTACGTAA 1840  
QY 61 ATCTGAGAGGTGACCAAGCTGACTGACCGTGGTACTCACATTTCTGCTGATCAGTC 120  
Db 1841 ATCTGAGAGGTGACCAAGCTGACTGACCGTGGTACTCACATTTCTGCTGATCAGTC 1900  
QY 121 AGCACTGTGACTCTTTTAAAGTATAGTAACAGAGAAATCCACATCTGTGTGCTCC 180  
Db 1901 AGCACTGTGACTCTTTTAAAGTATAGTAACAGAGAAATCCACATCTGTGTGCTCC 1960  
QY 181 TTAAAGAGCTATCAAGTCTGAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCA 240  
Db 1961 TTAAAGAGCTATCAAGTCTGAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCA 2020  
QY 241 CCATTTTGGATTGACAGGCTCTCTTCTGACTATACCTCTCCAGAAATTTTCTCTTTG 300  
Db 2021 CCATTTTGGATTGACAGGCTCTCTTCTGACTATACCTCTCCAGAAATTTTCTCTTTG 2080  
QY 301 AAAGTACTACTGGAATTTACATTTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTG 360  
Db 2081 AAAGTACTACTGGAATTTACATTTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTG 2140  
QY 361 GAAAGAAATATCCTACTGTGCTGTTTATATATGGTGGTCTCTC----- 402  
Db 2141 GAAAGAAATATCCTACTGTGCTGTTTATATATGGTGGTCTCTCAGGTGCGAGTTGGTAA 2200  
QY 403 ----- 402  
Db 2201 TAATCGGTTTAAAGGAGTCAAGTATTTTCGCTTGAATACCCCTAGCCTCTCTAGGTATGT 2260  
QY 403 ----- 402  
Db 2261 GGTGTAGTATAGACAACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAGCGCCTT 2320  
QY 403 -----AGGTCAAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGC 451  
Db 2321 TAAATATAAATGGGTCAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGC 2380  
QY 452 TTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGG 511  
Db 2381 TTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGG 2440  
QY 512 ATACCTCTCCCTGATGSCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATGCTGG 571  
Db 2441 ATACCTCTCCCTGATGSCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATGCTGG 2500  
QY 572 GGCCCCAGTCACTCTGTGATCTTCTATGATACAGATACACGGAAAGCTTATATGGGTCA 631  
Db 2501 GGCCCCAGTCACTCTGTGATCTTCTATGATACAGATACACGGAAAGCTTATATGGGTCA 2560  
QY 632 CCCTGACCAAGTGAACGCGCTATTACTTAGATCTGTGGGCATGCAAGCAGCAAAAGTT 691



```
Db 2921 TTTAATCAACAGAAAACACAGAAATTGATCATCACATTTTGTATACCTGCCATGTAAATCT 2980
Qy 1052 ACTCCTGAAATAAATGTGGTGCATGCGAGGGGTCTACGGTTTGTGGTGTAGTAACTTAATA 1111
Db 2981 ACTCCTGAAATAAATGTGGTGCATGCGAGGGGTCTACGGTTTGTGGTGTAGTAACTTAATA 3040
Qy 1112 CTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 1171
Db 3041 CTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3100
Qy 1172 TAAGAAATTACTAAAAAARAAAAA 1197
Db 3101 TAAGAAATTACTAAAAAARAAAAA 3126

RESULT 11
US-10-311-035-30
; Sequence 30; Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depodriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: FI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 3106
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CB1
US-10-311-035-30

Query Match 86.3%; Score 1033; DB 7; Length 3106;
Best Local Similarity 89.0%; Pred. No. 1.1e-300;
Matches 1190; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

Qy 1 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 60
Db 1770 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 1829
Qy 61 ATTCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTCTTCTGTCATCAGTC 120
```

```
Db 1830 ATCTCGAGAGGTGACAGAGCTGACTGACCGTGGCTACTCACTTTCTTGTGCATCAGTC 1889
Qy 121 AGCACTCTGACTTCTTTTATAGTAACTATAGTAAACAGAGAAATCCACACTGTGTGTCCC 180
Db 1890 AGCACTCTGACTTCTTTTATAGTAACTATAGTAAACAGAGAAATCCACACTGTGTGTCCC 1949
Qy 181 TTTTCAAGACTATCAAGTCTCTGAAGATGACCAACTTGGCAAAACAAAGAAATTTTGGGCCA 240
Db 1950 TTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTGGCAAAACAAAGAAATTTTGGGCCA 2009
Qy 241 CCATTTTGGATTGAGCAGGTCTCTTCTCTGACTATCTCTCTCCAGAAATTTTCTCTTTTG 300
Db 2010 CCATTTTGGATTGAGCAGGTCTCTTCTCTGACTATCTCTCTCCAGAAATTTTCTCTTTTG 2069
Qy 301 AAAGTACTACTGGAATTTACATTTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
Db 2070 AAAGTACTACTGGAATTTACATTTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 2129
Qy 361 GAAAGAAATATCCTACTGTGCTGTTTCAATATATGTTGGTCTCTCA----- 403
Db 2130 GAAAGAAATATCCTACTGTGCTGTTTCAATATATGTTGGTCTCTCAGGTGCAGTTGGTGAATA 2189
Qy 404 ----- 403
Db 2190 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAATACCTAGCCTCTCTAGTTATCTGG 2249
Qy 404 ----- 403
Db 2250 TTGTAGTGATAGACAACAGGGGATCCTGTCCACCGAGGGCTTAAATTTGAAGGGCCCTTTA 2309
Qy 404 -----GGTCAATAGAAATAGAGTACAGTACAGGTGAGAGGACTCGAATATCTAGCTT 453
Db 2310 AATATAAATGGGTCAATAGAAATGACATCAGTGGAAAGGACTCAATATCTAGCTT 2369
Qy 454 CTCGATATGATTTTCAATGATCTAGTCTGTGGGATCCACGGTGTCTCTATGGAGAT 513
Db 2370 CTCGATATGATTTTCAATGATCTAGTCTGTGGGATCCACGGTGTCTCTATGGAGAT 2429
Qy 514 ACCTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTCAGGGTGTCTATTTGCTGGG 573
Db 2430 ACCTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTCAGGGTGTCTATTTGCTGGG 2489
Qy 574 CCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACAGGAAACGTTATATGGGTCAAC 633
Db 2490 CCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACAGGAAACGTTATATGGGTCAAC 2549
Qy 634 CTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCC 693
Db 2550 CTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCC 2609
Qy 694 CCTCTGAACCAAAATCGTTTACTGTCTTACATGGTTTCTTGGATGAGAAATGTCCATTTTG 753
Db 2610 CCTCTGAACCAAAATCGTTTACTGTCTTACATGGTTTCTTGGATGAGAAATGTCCATTTTG 2669
Qy 754 CACATACCAAGTATATTTACTGAGTTTCTTGTAGGGCTGGAAAGCCATATGATTTACAGA 813
Db 2670 CACATACCAAGTATATTTACTGAGTTTCTTGTAGGGCTGGAAAGCCATATGATTTACAGA 2729
Qy 814 TCTATCCTCAGGAGAGACACAGCATAGAGTTCTGTAATCGGGAGACATTTAGAACTGC 873
Db 2730 TCTATCCTCAGGAGAGACACAGCATAGAGTTCTGTAATCGGGAGACATTTAGAACTGC 2789
Qy 874 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGTTATTTGCTGTAAAGTCAAT 933
Db 2790 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGTTATTTGCTGTAAAGTCAAT 2849
Qy 934 AATTTTGACCTGTGTAGAACTCTCTGTTATACACTGGCTATTTAACCAATAGAGAGTT 993
Db 2850 AATTTTGACCTGTGTAGAACTCTCTGTTATACACTGGCTATTTAACCAATAGAGAGTT 2909
Qy 994 TAACTCAAGAAACACAGAAATTCATCACATTTTGTATACCTGCGCATGTAACTCTAC 1053
```

Db 2910 TAAATCAACAGAAACACAGAAATTGATCATCATCATTTTGTATCTACCTGCCATGTAAACATCTAC 2969  
Qy 1054 TCTGAAATAAATGTGGTCCCATGACGGGGTCTACGGTTTGTGGTAGTAATCTAATACC 1113  
Db 2970 TCTGAAATAAATGTGGTCCCATGACGGGGTCTACGGTTTGTGGTAGTAATCTAATACC 3029  
Qy 1114 TTAACCCACATCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATA 1173  
Db 3030 TTAACCCACATCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATA 3089  
Qy 1174 AGAATTACTAAAAAAA 1190  
Db 3090 AGAATTACTAAAAAAA 3106

RESULT 12  
US-09-976-674-22  
; Sequence 22, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 4685  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-22

Query Match 84.7%; Score 1013.6; DB 3; Length 4685;  
Best Local Similarity 91.3%; Pred. No. 1e-294;  
Matches 1098; Conservative 0; Mismatches 99; Indels 5; Gaps 2;

Qy 1 ATTTTGAAGCACCAGAACTCCCTTTAGAGCATCACTGCTAGCTAGTCTAGTTACGTAA 60  
Db 1781 ATTTTGAAGCACCAGAACTCCCTTTAGAGCATCACTGCTAGCTAGTCTAGTTACGTAA 1840  
Qy 61 ATCTGAGAGGTGACAAAGGCTGACCTGACCGTGGCTACTCAATTTCTTGTGCTCAGTC 120  
Db 1841 ATCTGAGAGGTGACAAAGGCTGACCTGACCGTGGCTACTCAATTTCTTGTGCTCAGTC 1900  
Qy 121 AdCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACATCTGTGTGCC 180  
Db 1901 AdCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACATCTGTGTGCC 1960  
Qy 181 TTTCAGAGTATCAAGTCTTGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCA 240  
Db 1961 TTTCAGAGTATCAAGTCTTGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCA 2020  
Qy 241 CCATTTGGATTGAC-----CAGGTCTCTTCTGTGACT-ATACTCTCCAGAAATTTTCTC 295  
Db 2021 CCATTTGGATTGACCTCCTCAGGTGCACTTGGTGAATATATCGTTTAAAGAGTCAAGTA 2080  
Qy 296 TTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACA 355  
Db 2081 TTTCCGCTTGAATACCTCCTAGCTCTTAGGTATGTTGTTAGTATAGTACACACAGGGG 2140  
Qy 356 GCCTGGAAGAAATATCTACTGTCTGTCTATATATGTTGTTGCTCTCAGGCTCAATAGA 415  
Db 2141 ATCTGTGACCGAGGGCTTAAATTTGAGCGGCTTAAATATAAATGAGTCAATAGA 2200  
Qy 416 AATTGACGATCAGTGGAGGACTCCAAATATCTAGCTTCTCGATGATTTTCATGACTT 475  
Db 2201 AATTGACGATCAGTGGAGGACTCCAAATATCTAGCTTCTCGATGATTTTCATGACTT 2260

Qy 476 AGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAAT 535  
Db 2261 AGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAAT 2320  
Qy 536 GCAGAGGTCAAGATATCTTTCAGGGTGTCTATTTGCTGGGGCCCAAGTCACTCTGTGGATCTT 595  
Db 2321 GCAGAGGTCAAGATATCTTTCAGGGTGTCTATTTGCTGGGGCCCAAGTCACTCTGTGGATCTT 2380  
Qy 596 CTATGATACAGGATACACCGAACGTTATATGGGTCACTCCCTGACAGATGAACAGGGCTA 655  
Db 2381 CTATGATACAGGATACACCGAACGTTATATGGGTCACTCCCTGACAGATGAACAGGGCTA 2440  
Qy 656 TTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGTAACCAAAATCGTTTACT 715  
Db 2441 TTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGTAACCAAAATCGTTTACT 2500  
Qy 716 GCTCTTACATGGTTTCTCTGATGAGAAATGCCATTTTGGACATACCATGATATTTACTGAG 775  
Db 2501 GCTCTTACATGGTTTCTCTGATGAGAAATGCCATTTTGGACATACCATGATATTTACTGAG 2560  
Qy 776 TTTTATAGTGAAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAG 835  
Db 2561 TTTTATAGTGAAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAG 2620  
Qy 836 CATAGAGTTCTTGAATCGGAGAACATTTATGAATCTATCTATCTCTCAGGAGAGACACAG 895  
Db 2621 CATAGAGTTCTTGAATCGGAGAACATTTATGAATCTATCTATCTCTCAGGAGAGACACAG 2680  
Qy 896 AAACCTTGGATCAGTATTTGCTGCTCTAAAGTATATAATTTTACCTGTGTAGAACTC 955  
Db 2681 AAACCTTGGATCAGTATTTGCTGCTCTAAAGTATATAATTTTACCTGTGTAGAACTC 2740  
Qy 956 TCTGATATACATCTGGCTTATTTAAACCAATGAGAGGTTTAAATCAACAGAAAACACAGAAAT 1015  
Db 2741 TCTGATATACATCTGGCTTATTTAAACCAATGAGAGGTTTAAATCAACAGAAAACACAGAAAT 2800  
Qy 1016 TGATCATCATATTTGATACCTGCCATGTAACATCTACTCTCTGAAAAATAATGTGGTCCC 1075  
Db 2801 TGATCATCATATTTGATACCTGCCATGTAACATCTACTCTCTGAAAAATAATGTGGTCCC 2860  
Qy 1076 ATGAGGGGTCTACGGTTTGTGTAGTAATCTAAATCTTAAACCCACATGCTCAAAATC 1135  
Db 2861 ATGAGGGGTCTACGGTTTGTGTAGTAATCTAAATCTTAAACCCACATGCTCAAAATC 2920  
Qy 1136 AAATGATACATATTTCTGAGAGACCCAGAAATACCAATAGATTTACTAAAAAATAA 1195  
Db 2921 AAATGATACATATTTCTGAGAGACCCAGAAATACCAATAGATTTACTAAAAAATAA 2980  
Qy 1196 AA 1197  
Db 2981 AA 2982

RESULT 13  
US-10-982-512-22  
; Sequence 22, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; CURRENT FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-22

Query Match      84.7%; Score 1013.6; DB 9; Length 4685;
Best Local Similarity 91.3%; Pred. No. 18-294;
Matches 1098; Conservative 0; Mismatches 99; Indels 5; Gaps 2;

Qy 1 ATTTGAGGACCAAGAAGTCTCCCTTTAGAGCATCACTGACGTAGTACGTAA 60
Db ATTTGAGGACCAAGAAGTCTCCCTTTAGAGCATCACTGACGTAGTACGTAA 1840

Qy 61 ATCTGGAGAGTGACAAGGCTGACGTACCGTGGTACTCACTTTCTGTCATCAGTC 120
Db ATCTGGAGAGTGACAAGGCTGACGTACCGTGGTACTCACTTTCTGTCATCAGTC 1900

Qy 121 AGCACTGTGACTTTCTTTATAGTAAGTATAGTAACCAAGAATCCACATGTGTGCC 180
Db AGCACTGTGACTTTCTTTATAGTAAGTATAGTAACCAAGAATCCACATGTGTGCC 1960

Qy 181 TTTACAGCTATCAAGTCTGAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCA 240
Db TTTACAGCTATCAAGTCTGAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCA 2020

Qy 241 CCATTTGGATTTCAG---CAGTCTCTTCTGACT-ATACTCTCCAGAAATTTTCTC 295
Db CCATTTGGATTTCAGTCTCAGTCTGAGTGGTGAATATCGGTTTAAAGGATCAAGTA 2080

Qy 296 TTTTGAAGTACTACTGGATTACATTTGTGGATGCTCTACAAGCTCATGATCTACA 355
Db TTTCCGCTTGMATACCTAGCTCTCTAGTTATGTGGTTGTAGTATAGCAACAGGGG 2140

Qy 356 GCTCGGAAGAATATCTCTGCTGCTGTCATATATGGTGGTCTCAGGCTCAAAATAGA 415
Db ATCTGTCCAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAAATAGA 2200

Qy 416 AATTGACGATCAGTGGGAAGGATCCCAATATCTAGCTTCTCGATATGATTTCACTT 475
Db AATTGACGATCAGTGGGAAGGATCCCAATATCTAGCTTCTCGATATGATTTCACTT 2260

Qy 476 AGATCTGTGGGATCCACGGCTGCTCTATGGAAGTACTCTCTCCGTATGGCATTAAAT 535
Db AGATCTGTGGGATCCACGGCTGCTCTATGGAAGTACTCTCTCCGTATGGCATTAAAT 2320

Qy 536 GCAGAGTCAAGATATCTTACGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTT 595
Db GCAGAGTCAAGATATCTTACGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTT 2380

Qy 596 CTATGATACAGGATACAGGAACGTTATATGGGTCACTGACCAAGATGAACAGGGCTA 655
Db CTATGATACAGGATACAGGAACGTTATATGGGTCACTGACCAAGATGAACAGGGCTA 2440

Qy 656 TTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACT 715
Db TTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACT 2500

Qy 716 GCTCTTACATGTTTCTTGGATGAAATGTCATTTTGCATACCAAGTATATTACTGAG 775
Db GCTCTTACATGTTTCTTGGATGAAATGTCATTTTGCATACCAAGTATATTACTGAG 2560

Qy 776 TTTTGTAGTGGGCTGGAAGCCATATGATTTACATCTATTCCTCAGAGAGACAG 835
Db TTTTGTAGTGGGCTGGAAGCCATATGATTTACATCTATTCCTCAGAGAGACAG 2620

Qy 836 CATTAAGTCTCTGATCGGAGAACATATGAACTGATCTTTTGCATCTACCTCAAGA 895
Db CATTAAGTCTCTGATCGGAGAACATATGAACTGATCTTTTGCATCTACCTCAAGA 2680

Qy 896 AAACCTTGATCACTGATTTGCTCTTAAAGTGAATATAATTTTGACCTGTGTGAACCTC 955
Db AAACCTTGATCACTGATTTGCTCTTAAAGTGAATATAATTTTGACCTGTGTGAACCTC 2740
```

```
Qy 956 TCTGTATACACTGGCTATTTAACCAAAATGAGGAGGTTTAAATCAACAGAAAACACAGAAT 1015
Db 2741 TCTGTATACACTGGCTATTTAACCAAAATGAGGAGGTTTAAATCAACAGAAAACACAGAAT 2800

Qy 1016 TGATCATCACATTTTGTAGTACCTGCGCATGTAACATCTACTCTCTGAAAATAAAATGTGGTGCC 1075
Db 2801 TGATCATCACATTTTGTAGTACCTGCGCATGTAACATCTACTCTCTGAAAATAAAATGTGGTGCC 2860

Qy 1076 ATGAGGGGCTTACGGTTTGTGTAGTAAATCTTAATACCTTAAACCCCAATGCTCAAAATC 1135
Db 2861 ATGAGGGGCTTACGGTTTGTGTAGTAAATCTTAATACCTTAAACCCCAATGCTCAAAATC 2920

Qy 1136 AAATGATACATATTCCTGAGAGACCCAGCAATACCATTAAGAAATTACTAAAAAATAAAAAA 1195
Db 2921 AAATGATACATATTCCTGAGAGACCCAGCAATACCATTAAGAAATTACTAAAAAATAAAAAA 2980

Qy 1196 AA 1197
Db 2981 AA 2982

RESULT 14
US-10-825-632-6
; Sequence 6; Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABEOTI; Catherine Anne
; APPLICANT: GORRELL; Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1669
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-6
```

```
Query Match      73.6%; Score 881; DB 8; Length 1669;
Best Local Similarity 87.2%; Pred. No. 7,1e-255;
Matches 1044; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

Qy 1 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGTACGTAGTACGTAA 60
Db 618 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGTACGTAGTACGTAA 677

Qy 61 ATCTGGAGAGTGAACAAGGCTGACTGACCGTGGCTACTACATTTCTTGTGATCAATGTC 120
Db 678 ATCTGGAGAGTGAACAAGGCTGACTGACCGTGGCTACTACATTTCTTGTGATCAATGTC 737

Qy 121 AGCACTGTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180
Db 738 AGCACTGTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 797

Qy 181 TTTTCAAGCTATCAAGTCTTGAAGATGACCAACTTGCAAAACAAAGAAATTTTGGGCCA 240
Db 798 TTTTCAAGCTATCAAGTCTTGAAGATGACCAACTTGCAAAACAAAGAAATTTTGGGCCA 857

Qy 241 CCATTTTGGATTCAGCAGGCTCTCTTCTGACTATACCTCTCAGAAAATTTTCTTTTGG 300
Db 858 CCATTTTGGATTCAGCAGGCTCTCTTCTGACTATACCTCTCAGAAAATTTTCTTTTGG 917
```



```
QY 301 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
Db 918 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 977
QY 361 GAAAGAAATATCCTACTGCTGCTTTCATATATATGGTGGTCTCAGGGTCAAAATAGAAATG 420
Db 978 GAAAGAAATATCCTACTGCTGCTTTCATATATATGGTGGTCTCA----- 1020
QY 421 ACGATCAGGTGGAGAGCTCAATATATCTAGCTTCTCGATATGATTTCAATGACTAGATC 480
Db 1021 ----- 1020
QY 481 GTGTGGGCATCCACGGCTGGTCTATGGAGGATACCTCTCCCTGATGCAATTAATGCAGA 540
Db 1021 ----- 1020
QY 541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATG 600
Db 1021 -----GTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATG 1064
QY 601 ATACAGGATACAGGAACGTTATATGGGTCAACCTGACCAAGTAAACAGGGCTATTACT 660
Db 1065 ATACAGGATACAGGAACGTTATATGGGTCAACCTGACCAAGTAAACAGGGCTATTACT 1124
QY 661 TAGGATCTGTGGGCATGCAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 720
Db 1125 TAGGATCTGTGGGCATGCAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 1184
QY 721 TACATGTTTCTCGGATGAGAAATGTCATTTTGCACATACCAAGTATATTACTAGTATTT 780
Db 1185 TACATGTTTCTCGGATGAGAAATGTCATTTTGCACATACCAAGTATATTACTAGTATTT 1244
QY 781 TAGTGAGGCTGGAGAGCCATATGATTTACAGATCTATCTCCAGAGAGACACAGCATAA 840
Db 1245 TAGTGAGGCTGGAGAGCCATATGATTTACAGATCTATCTCCAGAGAGACACAGCATAA 1304
QY 841 GAGTTCCTGTAATCGGAGAACATTATGAATCTGATCTTTTGACCTTACCTTCAAGAAACC 900
Db 1305 GAGTTCCTGTAATCGGAGAACATTATGAATCTGATCTTTTGACCTTACCTTCAAGAAACC 1364
QY 901 TTGGATCACGTATTGCTGCTCTAAAAGTATATATTTTGACCTGTGTAGAACCTCTG 960
Db 1365 TTGGATCACGTATTGCTGCTCTAAAAGTATATATTTTGACCTGTGTAGAACCTCTG 1424
QY 961 TATACATGCTATTATTAACCAATGAGAGGTTTATATCAACAGAAACACAGAAATGATC 1020
Db 1425 TATACATGCTATTATTAACCAATGAGAGGTTTATATCAACAGAAACACAGAAATGATC 1484
QY 1021 ATACATTTTGATACCTGCCATGTAACATCTACTCTGAAATTAATGTTGGTGCATGCA 1080
Db 1485 ATACATTTTGATACCTGCCATGTAACATCTACTCTGAAATTAATGTTGGTGCATGCA 1544
QY 1081 GGGGCTACGGTTTGTGATGATTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATG 1140
Db 1545 GGGGCTACGGTTTGTGATGATTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATG 1604
QY 1141 ATACATATTTCTGAGAGACCCAGCAATACCAATAGAAATTTACTTAAAAAATAAAAA 1197
Db 1605 ATACATATTTCTGAGAGACCCAGCAATACCAATAGAAATTTACTTAAAAAATAAAAA 1661
```

RESULT 15  
US-10-956-157-2177  
; Sequence 2177, Application US/10956157  
; Publication No. US20050118625A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William  
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH  
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES  
; FILE REFERENCE: 031896-043000 (AM 101081)  
; CURRENT APPLICATION NUMBER: US/10/956,157  
; CURRENT FILING DATE: 2004-10-04

```
; NUMBER OF SEQ ID NOS: 319805  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2177  
; LENGTH: 2830  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-956-157-2177  
  
Query Match 73.6%; Score 881; DB 9; Length 2830;  
Best Local Similarity 87.2%; Pred. No; 9.7e-255;  
Matches 1044; Conservative 0; Mismatches 0; Indels 153; Gaps 1;  
  
QY 1 ATTTTGAAGCAGCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGCTAGTACGTAA 60  
Db 1781 ATTTTGAAGCAGCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 1840  
QY 61 ATCTCTGAGAGGTGCAAGGCTGACTGAGCGTGGGTACTCACATCTTCTGTGCAATCAGTC 120  
Db 1841 ATCTCTGAGAGGTGCAAGGCTGACTGAGCGTGGGTACTCACATCTTCTGTGCAATCAGTC 1900  
QY 121 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTC 180  
Db 1901 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTC 1960  
QY 181 TTTTAAAGCTATCAAGTCTGAAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGGCA 240  
Db 1961 TTTTAAAGCTATCAAGTCTGAAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGGCA 2020  
QY 241 CCATTTTGGGATTCAGCAGGTCCTTCTGACTATCTCTCCAGAAATTTTCTCTTTG 300  
Db 2021 CCATTTTGGGATTCAGCAGGTCCTTCTGACTATCTCTCCAGAAATTTTCTCTTTG 2080  
QY 301 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360  
Db 2081 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 2140  
QY 361 GAAAGAAATATCCTACTGCTCTTTCATATATGGTGGTCTCAGGGTCAAAATAGAAATG 420  
Db 2141 GAAAGAAATATCCTACTGCTCTTTCATATATGGTGGTCTCA----- 2183  
QY 421 ACGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATTTAGTATC 480  
Db 2184 ----- 2183  
QY 481 GTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGCAATTAATGCAGA 540  
Db 2184 ----- 2183  
QY 541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATG 600  
Db 2184 -----GTTTGTCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATG 2227  
QY 601 ATACAGGATACAGGAACGTTATATGGGTCAACCTGACCAAGTAAACAGAGGCTATTACT 660  
Db 2228 ATACAGGATACAGGAACGTTATATGGGTCAACCTGACCAAGTAAACAGAGGCTATTACT 2287  
QY 661 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 720  
Db 2288 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 2347  
QY 721 TACATGTTTCTCGGATGAGAAATGTCATTTTGCACATACCAAGTATATTACTAGTATTTT 780  
Db 2348 TACATGTTTCTCGGATGAGAAATGTCATTTTGCACATACCAAGTATATTACTAGTATTTT 2407  
QY 781 TAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAA 840  
Db 2408 TAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAA 2467  
QY 841 GAGTTCCTGTAATCGGAGAACATTATGAATCTCTTTTGCACCTTACCTTCAAGAAACC 900  
Db 2468 GAGTTCCTGTAATCGGAGAACATTATGAATCTCTTTTGCACCTTACCTTCAAGAAACC 2527  
QY 901 TTGGATCACGTATTGCTGCTCTAAAAGTATATATTTTGACCTGTGTAGAACTCTCTGG 960
```



Mon May 8 13:41:55 2006

Db	2528	TTGGATACGATTTCCTGCTCTAAAGTGATATAATTTTGACCTGTGAGAACTCTCTGG	2587
Qy	961	TATACACTGGCTATTATTAACCAATGAGGAGGTTTAATCAACAGAAAAACAGAAATTCGATC	1020
Db	2588	TATACACTGGCTATTATTAACCAATGAGGAGGTTTAATCAACAGAAAAACAGAAATTCGATC	2647
Qy	1021	ATCACATTTTGATACCTGCGCATGTAAACATCTACTCTCGAAATAAATGTGTGCGCATGCA	1080
Db	2648	ATCACATTTTGATACCTGCGCATGTAAACATCTACTCTGAAATAAATGTGTGCGCATGCA	2707
Qy	1081	GGGGTCACGGTTTGTGGTAGTAATCTAATACCTTAACCCCATGCTCAAAATCAAATG	1140
Db	2708	GGGGTCACGGTTTGTGGTAGTAATCTAATACCTTAACCCCATGCTCAAAATCAAATG	2767
Qy	1141	ATACATATTCCTGAGAGACCCGACCAATACCATAGAATTTACTTAAAAAATAAAAAAAA	1197
Db	2768	ATACATATTCCTGAGAGACCCGACCAATACCATAGAATTTACTTAAAAAATAAAAAAAA	2824

Search completed: May 4, 2006, 11:21:58  
Job time : 1070.84 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:53:29 ; Search time 5966.21 Seconds  
(without alignments)  
11404.501 Million cell updates/sec

Title: US-10-825-632-4  
Perfect score: 1197  
Sequence: 1 attttgaagcaccacaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pt.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1197	100.0	1197	6 AR651453	AR651453 Sequence
2	1197	100.0	1197	8 AF221635	AF221635 Homo sapi
3	1179	98.5	4676	6 AR631287	AR631287 Sequence
4	1179	98.5	4676	6 AX608743	AX608743 Sequence
5	1040	86.9	2797	6 AR448400	AR448400 Sequence
6	1040	86.9	3127	8 AF221634	AF221634 Homo sapi
7	1040	86.9	3143	6 AX354793	AX354793 Sequence
8	1036	86.5	3120	6 AR651452	AR651452 Sequence
9	1036	86.5	4535	8 BC040203	BC040203 Homo sapi
10	1033.2	86.3	4309	6 AR631284	AR631284 Sequence
11	1033.2	86.3	4309	6 AX608737	AX608737 Sequence
12	1033.2	86.3	4829	6 AR631283	AR631283 Sequence
13	1033.2	86.3	4829	6 AX608735	AX608735 Sequence
14	1033	86.3	3106	6 AX342633	AX342633 Sequence
15	1016.2	84.9	3125	8 BC030688	BC030688 Homo sapi
16	1013.6	84.7	4685	6 AR631288	AR631288 Sequence
17	1013.6	84.7	4685	6 AX608745	AX608745 Sequence
18	999.6	83.5	2161	6 BD157001	BD157001 Primer fo

19	999.6	83.5	2161	6 AX878058	AX878058 Sequence
20	999.6	83.5	2161	8 AK027826	AK027826 Homo sapi
21	923	77.1	3030	8 AY354202	AY354202 Homo sapi
22	881	73.6	1669	6 AR651454	AR651454 Sequence
23	881	73.6	1669	8 AF221636	AF221636 Homo sapi
24	881	73.6	2778	8 AK000290	AK000290 Homo sapi
25	879	73.4	2668	6 AX405771	AX405771 Sequence
26	879	73.4	2842	6 AX405770	AX405770 Sequence
27	873	72.9	4523	6 AR631281	AR631281 Sequence
28	873	72.9	4523	6 AX608731	AX608731 Sequence
29	863.6	72.1	2510	6 AX338497	AX338497 Sequence
30	791.4	66.1	873	6 AX083130	AX083130 Sequence
31	778	65.0	2649	6 AX354795	AX354795 Sequence
32	776.4	64.9	2656	8 AF172659	AF172659 Homo sapi
33	776.4	64.9	2671	6 AR631278	AR631278 Sequence
34	776.4	64.9	2671	6 AX608725	AX608725 Sequence
35	704.8	58.9	4799	9 BC043124	BC043124 Mus muscu
36	704.8	58.9	4799	9 BC059222	BC059222 Mus muscu
37	677	56.6	735	6 AX524942	AX524942 Sequence
38	633.4	52.9	823	6 AR263993	AR263993 Sequence
39	599.2	50.1	1002	6 AX524939	AX524939 Sequence
40	540	45.1	2098	6 CQ721272	CQ721272 Sequence
41	523.6	43.7	587	6 BD154822	BD154822 Primer fo
42	523.6	43.7	587	6 AX874760	AX874760 Sequence
43	500.6	41.8	1449	5 CR353491	CR353491 Gallus ga
44	459	38.3	1083	6 AR651455	AR651455 Sequence
45	459	38.3	1083	8 AF221637	AF221637 Homo sapi

ALIGNMENTS

RESULT 1  
AR651453  
LOCUS AR651453 1197 bp DNA linear PAT 20-APR-2005  
DEFINITION Sequence 4 from patent US 6881564.  
ACCESSION AR651453  
VERSION AR651453.1 GI:62795939  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1197)  
AUTHORS Abbott,C.A. and Gorrell,M.D.  
TITLE Dipeptidyl peptidases  
JOURNAL Patent: US 6881564-A 4 19-APR-2005;  
The University of Sydney; Sydney;  
AUX;  
FEATURES  
source Location/Qualifiers  
1..1197  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN	Query Match	Best Local Similarity	100.0%;	Score 1197;	DB 6;	Length 1197;
	Matches 1197;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA	60			
Db	1	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA	60			
Qy	61	ATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTCTGCATCAGTC	120			
Db	61	ATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTCTGCATCAGTC	120			
Qy	121	AGCACTGTGACTTCTTTATAGTAAGTATAGTAACACAGAATCCACACTGTGTGTCC	180			
Db	121	AGCACTGTGACTTCTTTATAGTAAGTATAGTAACACAGAATCCACACTGTGTGTCC	180			
Qy	181	TTTACAGCTATCAAGTCTTCTTCTGAGAGTACCCCACTTGCAGAAAGGAATTTTGGGCCA	240			
Db	181	TTTACAGCTATCAAGTCTTCTTCTGAGAGTACCCCACTTGCAGAAAGGAATTTTGGGCCA	240			

QY	241	CCATTTTGGATTGACGAGGTCCTTCTTCCTGACTATCTCTCCAGAAATTTTCTCTTTTG	300
Db	241	CCATTTTGGATTGACGAGGTCCTTCTTCCTGACTATCTCTCCAGAAATTTTCTCTTTTG	300
QY	301	AAAGTACTACTGGAATTAATTTGATGGAGTCTCTACAGGCTCATGATCTACAGCTG	360
Db	301	AAAGTACTACTGGAATTAATTTGATGGAGTCTCTACAGGCTCATGATCTACAGCTG	360
QY	361	GAAGAATAATCTCTACTGCTCTTATATATGGTGGTCTCAGGGTCAAAATAGAAATG	420
Db	361	GAAGAATAATCTCTACTGCTCTTATATATGGTGGTCTCAGGGTCAAAATAGAAATG	420
QY	421	ACGATCAGGTGGAGGCTCCAAATATCTAGCTTCTCGATGATTTTCAATGACTTAGATC	480
Db	421	ACGATCAGGTGGAGGCTCCAAATATCTAGCTTCTCGATGATTTTCAATGACTTAGATC	480
QY	481	GTGTGGGCTACCCAGGCTGCTCTATGGAGGATACCTCTCCCTGATGCATTAATGAGA	540
Db	481	GTGTGGGCTACCCAGGCTGCTCTATGGAGGATACCTCTCCCTGATGCATTAATGAGA	540
QY	541	GGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCATG	600
Db	541	GGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCATG	600
QY	601	ATACAGGATACACGGAACGTTATATGGGTCACTCTGACAGAAATGAAACAGGGCTATTACT	660
Db	601	ATACAGGATACACGGAACGTTATATGGGTCACTCTGACAGAAATGAAACAGGGCTATTACT	660
QY	661	TAGGATCTGTGGGCTACGAGCAAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT	720
Db	661	TAGGATCTGTGGGCTACGAGCAAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT	720
QY	721	TACATGCTTCTCGATGAGAAATGTCATTTTGCAATACCAAGTATATCTAGGTTTTT	780
Db	721	TACATGCTTCTCGATGAGAAATGTCATTTTGCAATACCAAGTATATCTAGGTTTTT	780
QY	781	TAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGAGGAGACAGCAATAA	840
Db	781	TAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGAGGAGACAGCAATAA	840
QY	841	GAGTCTCTGAATCGGGAACATATTAATGAACTGATCTTTTGCACTACCTTCAAGAAACC	900
Db	841	GAGTCTCTGAATCGGGAACATATTAATGAACTGATCTTTTGCACTACCTTCAAGAAACC	900
QY	901	TTGGATCAGATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG	960
Db	901	TTGGATCAGATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG	960
QY	961	TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATGATC	1020
Db	961	TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATGATC	1020
QY	1021	ATCACATTTTGATACCTGCCATGTAAATCTACTCTCTGAAATAAATGTGTGCCATGCA	1080
Db	1021	ATCACATTTTGATACCTGCCATGTAAATCTACTCTCTGAAATAAATGTGTGCCATGCA	1080
QY	1081	GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATG	1140
Db	1081	GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATG	1140
QY	1141	ATACATATTTCTGAGAGCCAGCAATACATAAGAAATTTACTTAAAAAATTTTTTTTTT	1197
Db	1141	ATACATATTTCTGAGAGCCAGCAATACATAAGAAATTTACTTAAAAAATTTTTTTTTT	1197
RESULT 2			
AF221635			
LOCUS			
DEFINITION Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, partial cds,			
alternatively spliced.			
AF221635			
ACCESSION			
VERSION AF221635.1 GI:11095189			
KEYWORDS			
SOURCE Homo sapiens (human)			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1..1197			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="15"			
/map="15q22"			
/clone="721"			
/tissue_type="placenta"			
/clone_lib="lambdaTriplexExTM: Clontech"			
<1..1197			
/gene="DPP8"			
<1..935			
/genes="DPP8"			
/note="serine protease; member of prolyl oligopeptidase S9b			
family belonging to enzyme clan SC; additional splice			
variants of this mRNA are found in GenBank Accession			
Numbers AF221634, AF221636 and AF221637"			
/codon_start=3			
/product="dipeptidyl peptidase 8"			
/protein_id="AAG29767.1"			
/db_xref="GI:11095190"			
/translation="PEGTVDSPLEHLVVSVPGEVTRLTDRGSHSCISQHCDF			
FISKSNQNPCHVSLYKLSDEDDTCKTREFWATLDSAGPLPDTPPEIPSEFT			
TFTLGMLYKPEDDQPKKXTVLFTVGQGIIEIDVDGEGLOVLYASHYDFDLDR			
YHGHSYGYGLSLMALNQRSDIFVAAGAPVTLMEFYDGYTERVYNGHPDNEQY			
YLSGVAMQAEKFPSEPNRLLLHGFIDENVFAHTSILLSLFLVRAGKPYDLQIYPQR			
HSRVPESGEHVELHLLHYLQENLGSRAALKVI"			
ORIGIN			
Query Match 100.0%; Score 1197; DB 8; Length 1197;			
Best Local Similarity 100.0%; Pred. No. 1.2e-308; Indels 0; Gaps 0;			
Matches 1197; Conservative 0; Mismatches 0;			
QY	1	ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTACGTACGTACGTACGTA	60
Db	1	ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTACGTACGTACGTACGTA	60
QY	61	ATCTGGAGAGGTGACAAGCTGACCTGGCTGCTACATCTCTTCTGCTGCAATGTC	120
Db	61	ATCTGGAGAGGTGACAAGCTGACCTGGCTGCTACATCTCTTCTGCTGCAATGTC	120
QY	121	AGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAATCCACACTGTGTGCTCC	180
Db	121	AGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAATCCACACTGTGTGCTCC	180
QY	181	TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACCAAGAAATTTTGGGCA	240
Db	181	TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACCAAGAAATTTTGGGCA	240
QY	241	CCATTTTGGATTGACGAGGTCCTTCTTCCTGACTATCTCTCCAGAAATTTTCTCTTTG	300
Db	241	CCATTTTGGATTGACGAGGTCCTTCTTCCTGACTATCTCTCCAGAAATTTTCTCTTTG	300



```
QY 901 TTGGATCAGCTATTGCTCTTAAGCTGATATATTTTGACCTGTGTAGAACTCTCTGG 960
Db 2677 TTGGATCAGCTATTGCTCTTAAGCTGATATATTTTGACCTGTGTAGAACTCTCTGG 2736
QY 961 TATACACTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTGATC 1020
Db 2737 TATACACTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTGATC 2796
QY 1021 ATACATATTTGATACCTGGCATGTAAACATCTACTCTGAAATAAATGTGTGCCATGCA 1080
Db 2797 ATACATATTTGATACCTGGCATGTAAACATCTACTCTGAAATAAATGTGTGCCATGCA 2856
QY 1081 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAAATCAAAATG 1140
Db 2857 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAAATCAAAATG 2916
QY 1141 ATACATATTTCTGAGAGACCCAGCAATACATTAAGAAATTTACTTAAAAAATAAAAAA 1197
Db 2917 ATACATATTTCTGAGAGACCCAGCAATACATTAAGAAATTTACTTAAAAAATAAAAAA 2973

RESULT 4
AX608743
LOCUS AX608743 4676 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 20 from Patent WO0231134.
ACCESSION AX608743
VERSION AX608743.1 GI:28404307
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
AUTHORS Qi,S., Akimanya,K.O., Riviere,P.J. and Junien,J.L.
TITLE Novel serine protease genes related to dppv
JOURNAL Patent: WO 0231134-A 20 18-APR-2002;
Ferring BV (NL)
FEATURES
    source          Location/Qualifiers
    1..4676
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Query Match 98.5%; Score 1179; DB 6; Length 4676;
Best Local Similarity 99.7%; Pred. No. 6.7e-304;
Matches 1193; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGTAGTCAGTTAGGTAA 60
Db 1781 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGTAGTCAGTTAGGTAA 1840
QY 61 ATCTGGAGGAGTGACNAGGCTGACTGACCGTGGTACTGACATTTCTGCTGCATCAGTC 120
Db 1841 ATCTGGAGGAGTGACNAGGCTGACTGACCGTGGTACTGACATTTCTGCTGCATCAGTC 1900
QY 121 AGCACTGTGACTCTTTATAGTATAGTATAGTAACAGAGAAATCCACATGTGTGTC 180
Db 1901 AGCACTGTGACTCTTTATAGTATAGTATAGTAACAGAGAAATCCACATGTGTGTC 1960
QY 181 TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAACAAAGAAATTTGGGCCA 240
Db 1961 TTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAACAAAGAAATTTGGGCCA 2020
QY 241 CCATTTTGGATTACAGAGTCTCTCTCTGACTATACCTCTCCAGAAATTTCTCTTTTG 300
Db 2021 CCATTTTGGATTACAGAGTCTCTCTCTGACTATACCTCTCCAGAAATTTCTCTTTTG 2080
QY 301 AAAGTACTACTGGATTACATTGTATGGATGCTCTCAAGCCCTCATGATCTACAGCCCTG 360
Db 2081 AAAGTACTACTGGATTACATTGTATGGATGCTCTCAAGCCCTCATGATCTACAGCCCTG 2140
```

```
QY 361 GAAAGAAATATCTCTACTGTGCTTTCATATATATGTTGGTCTCTCAGGGTCAAAATAGAAATG 420
Db 2141 GAAAGAAATATCTCTACTGTGCTTTCATATATATGTTGGTCTCTCAGGGTCAAAATAGAAATG 2196
QY 421 ACGATCAGGTGGAGGAGCCTCAATATCTAGCTTCTCGATATATGATTTTCACTTAGATC 480
Db 2197 ACGATCAGGTGGAGGAGCCTCAATATCTAGCTTCTCGATATATGATTTTCACTTAGATC 2256
QY 481 GTGTGGGCATCCACGGCTGGTCTCTATGAGGAGTACCTCTCCCTGTATGTCATTAATGAGA 540
Db 2257 GTGTGGGCATCCACGGCTGGTCTCTATGAGGAGTACCTCTCCCTGTATGTCATTAATGAGA 2316
QY 541 GGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATG 600
Db 2317 GGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATG 2376
QY 601 ATACAGGATACACGGAACTTATATGGGTACCCCTGACCAAGTGAACAGGGCTATTACT 660
Db 2377 ATACAGGATACACGGAACTTATATGGGTACCCCTGACCAAGTGAACAGGGCTATTACT 2436
QY 661 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720
Db 2437 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 2496
QY 721 TACATGGTTTCTCTGGATGAGAAATGTCATTTTGGCATACACAGTATATATTACTAGTTTT 780
Db 2497 TACATGGTTTCTCTGGATGAGAAATGTCATTTTGGCATACACAGTATATATTACTAGTTTT 2556
QY 781 TAGTGGGCTGAAAGCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATAA 840
Db 2557 TAGTGGGCTGAAAGCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATAA 2616
QY 841 GAGTTCCTGAATCGGAGAACATTATGAATGCTCATCTTTTGCATCTTCAAGAAAACC 900
Db 2617 GAGTTCCTGAATCGGAGAACATTATGAATGCTCATCTTTTGCATCTTCAAGAAAACC 2676
QY 901 TTGGATCACGTATTGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG 960
Db 2677 TTGGATCACGTATTGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG 2736
QY 961 TATACACTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTCATC 1020
Db 2737 TATACACTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAAATTCATC 2796
QY 1021 ATCACATTTTGTATACCTGCTGATGTAACATCTACTCTGAAAATAAATGTTGGTGCATGCA 1080
Db 2797 ATCACATTTTGTATACCTGCTGATGTAACATCTACTCTGAAAATAAATGTTGGTGCATGCA 2856
QY 1081 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAAATCAAAATG 1140
Db 2857 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAAATCAAAATG 2916
QY 1141 ATACATATTTCTGAGAGACCCAGCAATACCAATGAAATTTACTTAAAAAATAAAAAA 1197
Db 2917 ATACATATTTCTGAGAGACCCAGCAATACCAATGAAATTTACTTAAAAAATAAAAAA 2973

RESULT 5
AX448400
LOCUS AX448400 2797 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1103 from patent US 6673549.
ACCESSION AX448400
VERSION AX448400.1 GI:42676724
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2797)
AUTHORS Furness,L.M. and Buchbinder,J.L.
TITLE Genes expressed in C3A liver cell cultures treated with steroids
JOURNAL Patent: US 6673549-A 1103 06-JAN-2004;
Incyte Corporation; Palo Alto, CA
```

```
FEATURES
  source      Location/Qualifiers
 1..2797
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
Query Match      86.9%; Score 1040; DB 6; Length 2797;
Best Local Similarity 89.1%; Pred. No. 9.8e-267;
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 AATTTGAAGGCCAACAAAGACTCCCTTTAGAGCATACCTGTAGCTAGTACAGTAA 60
DB 1452 AATTTGAAGGCCAACAAAGACTCCCTTTAGAGCATACCTGTAGCTAGTACAGTAA 1511
QY 61 ATCTCTGGAGGTGACAGAGCTGACTGACCTGTGCTACTACATTCTTGTGTCATCAGTC 120
DB 1512 ATCTCTGGAGGTGACAGAGCTGACTGACCTGTGCTACTACATTCTTGTGTCATCAGTC 1571
QY 121 AGCACTGTGACTCTTTTAACTAAGTATAGTAAACCAAGAAATCCACACTGTGTCTCC 180
DB 1572 AGCACTGTGACTCTTTTAACTAAGTATAGTAAACCAAGAAATCCACACTGTGTCTCC 1631
QY 181 TTACAAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGAAATTTGGGCCA 240
DB 1632 TTACAAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGAAATTTGGGCCA 1691
QY 241 CCATTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 300
DB 1692 CCATTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTTG 1751
QY 301 AAGTACTACTGATTTATCATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTG 360
DB 1752 AAGTACTACTGATTTATCATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTG 1811
QY 361 GAAAGAAATATCTACTGTCTGTTTCATATATGTTGGTCTCA ----- 403
DB 1812 GAAAGAAATATCTACTGTCTGTTTCATATATGTTGGTCTCTCAGGTGAGTTGGTGAATA 1871
QY 404 ----- 403
DB 1872 ATCCGTTTAAAGGAGTCAAGTATTTCCGTTGAATACCTAGCTCTCTAGGTATGTGG 1931
QY 404 ----- 403
DB 1932 TTGTAGTGATAGACACAGGGGATCTCTGACCGAGGCTTAATTTGAAGGCCCTTTA 1991
QY 404 -----GGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCAATATCTAGCTT 453
DB 1992 AATATAAATGGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCAATATCTAGCTT 2051
QY 454 CTGATATGATTTCACTGATCTAGATCTGTTGGGATCCACGGCTGCTCTATGGAGGAT 513
DB 2052 CTGATATGATTTCACTGATCTAGATCTGTTGGGATCCACGGCTGCTCTATGGAGGAT 2111
QY 514 ACCTCTCCCTGATGGCAATTAATGACAGGTGAGATATCTCAGGGTTCCTATTCCTGGGG 573
DB 2112 ACCTCTCCCTGATGGCAATTAATGACAGGTGAGATATCTCAGGGTTCCTATTCCTGGGG 2171
QY 574 CCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC 633
DB 2172 CCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC 2231
QY 634 CTGACCAAGATGAACAGGGCTATTTACTTATGATCTGTGGCCATGCAAGCAAGAAAGTCC 693
DB 2232 CTGACCAAGATGAACAGGGCTATTTACTTATGATCTGTGGCCATGCAAGCAAGAAAGTCC 2291
QY 694 CCTCTGACCAAAATCGTTTACTCTCTTATCATGTTTCTCTGATGAGAAATGTCATTTTG 753
DB 2292 CCTCTGACCAAAATCGTTTACTCTCTTATCATGTTTCTCTGATGAGAAATGTCATTTTG 2351
QY 754 CACATACCAGTATATTAAGTATGAGTCTGAAAGCCATATGATTTACAGA 813
DB 2352 CACATACCAGTATATTAAGTATGAGTCTGAAAGCCATATGATTTACAGA 2411
```

```
QY 814 TCTATCTCTCAGGAGACACAGCATAGAGTTCCTGAATCGGAGAGACATTATGAACCTGC 873
DB 2412 TCTATCTCTCAGGAGACACAGCATAGAGTTCCTGAATCGGAGAGACATTATGAACCTGC 2471
QY 874 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCACTGATTCCTCTTAAAGTGTATAT 933
DB 2472 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCACTGATTCCTCTTAAAGTGTATAT 2531
QY 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAATGAGGAGTT 993
DB 2532 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAATGAGGAGTT 2591
QY 994 TAATCAACAGAAACACAGAAATTCATCATCACTTTTGATACCTGCCATGTAACATCTAC 1053
DB 2592 TAATCAACAGAAACACAGAAATTCATCATCACTTTTGATACCTGCCATGTAACATCTAC 2651
QY 1054 TCTGAAAAATAATGTGGTGCCTATCGAGGCTCTACGGTTTGTGTAGTATCTTAATACC 1113
DB 2652 TCTGAAAAATAATGTGGTGCCTATCGAGGCTCTACGGTTTGTGTAGTATCTTAATACC 2711
QY 1114 TTAACCCCATGCTCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCATA 1173
DB 2712 TTAACCCCATGCTCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCATA 2771
QY 1174 AGAATTACTAAAAAATAAAAAA 1197
DB 2772 AGAATTACTAAAAAATAAAAAA 2795

RESULT 6
AF221634 3127 bp mRNA linear PRI 05-NOV-2000
LOCUS Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, complete cds.
DEFINITION
ACCESSION AF221634
VERSION AF221634.1 GI:11095187
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 3127)
AUTHORS Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.
and Gorrell,M.D.
TITLE Cloning, expression and chromosomal localization of a novel human
dipeptidyl peptidase (DPP) IV homolog, DPP8
JOURNAL Eur. J. Biochem. 267 (20), 6140-6150 (2000)
PUBMED 11012666
AUTHORS Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver
Centre, Centenary Institute of Cell Biology and Cancer Medicine,
Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia
FEATURES
  source
 1..3127
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15q22"
 1..3127
 /gene="DPP8"
 214..2862
 /gene="DPP8"
 /note="serine protease; member of prolyl oligopeptidase S9b
 family belonging to enzyme clan SC; alternatively spliced
 forms of this mRNA are found in AF221635, AF221636 and
 AF221637"
 /codon_start=1
 /product="dipeptidyl peptidase 8"
 /protein_id="AAG29766.1"
```

/db\_xref="GI:11095188"  
/translations="MAAAMEIQGLVETADCEENIESODRPLKPFVVERYSWSQ  
LKKLLADTRKVGHYMAKAPDPMFKVGRNDPDGPHSDRIYVILAMSGENRNTLFTSEI  
PKTINRAAVLMLSKPKLLOLFOALDQVGMYSREBELAERXIGTIGASIDYHQSG  
TELFQAGSGTYHVKDGPQGFQOPLBNLJETSCHTRMDPKUCPADPWIAFIHNS  
DIWISNIVTFRERLJVAVNELANMEEDASAGUATFVJQEFDRISGYMWCPEAET  
PSGGKILILYEENDESSEVEI THVTSPLSTRADSFRPKGTANPKVTFKMSIEMI  
DAGRLIDIVDELIOPRILFEQVEYANAGWTFEGYAWISILLDRSQRLQIVLIS  
PELPIVEDDVMERGLIESVDSPLIITETIDMINIHDIFHVPQSHBEERIEF  
IPASECKTGRHLXYKITS LKESKYRSRSGGULPDSDFKPIKEEIAITSGEWELGR  
HGSNIQVDEVRILVPEGTRDSPLERHLIYVSNVPEVTKLDRGSHSCCSQHCED  
FRYSKSNKPNCHSLYKLSFDEDDTCKTKEFWATILDSAGPLPYTPPEIFSES  
TGFTLYGCKYRPHDLQPKQKPYTLVFIYGGQVOLVNNRPGVKVPRLNTLASLOY  
VVUIDNRGCKYRPHDLQPKQKPYTLVFIYGGQVOLVNNRPGVKVPRLNTLASLOY  
CGYLSMALMQSLDFRVAIAGAPVILMIFDYDTGYRMYGHPDQNGYLGVSAMQ  
AERFPEPRLHLLGDFDENVHFATISILLSFLVRAGKPYDLQIYPOERHSIRVPES  
GEHYELHLHYLOENLGSRIALKVI"

ORIGIN

Query Match	86.9%; Score 1040; DB 8; Length 3127;
Best Local Similarity	89.1%; Pred. No. 9.7e-267;
Matches 1197; Conservative	0; Mismatches 0; Indels 147; Gaps 1;
Qy	1 ATTTGAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAA 60
Db	1781 ATTTGAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAA 1840
Qy	61 ATCTGGAGGATGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTC 120
Db	1841 ATCTGGAGGATGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTC 1900
Qy	121 AGCACTGTGACTCTTTTAAAGTAAAGTATAGTACCAAGAAATCCACACTGTGTCCC 180
Db	1901 AGCACTGTGACTCTTTTAAAGTAAAGTATAGTACCAAGAAATCCACACTGTGTCCC 1960
Qy	181 TTTCACAGCTATCAAGTCTGAAGATGACCAACTTGCACAAACAAAGGATTTTGGCCA 240
Db	1961 TTTCACAGCTATCAAGTCTGAAGATGACCAACTTGCACAAACAAAGGATTTTGGCCA 2020
Qy	241 CCATTTTGGATTCAGCAGTCTCTTCTGCTGACTATCTCTCCAGAAATTTCTCTTTG 300
Db	2021 CCATTTTGGATTCAGCAGTCTCTTCTGCTGACTATCTCTCCAGAAATTTCTCTTTG 2080
Qy	301 AAGTACTACTGATTTACATTTGATGGATGCTCTACRAGCCTCATGTACTACAGCTG 360
Db	2081 AAGTACTACTGATTTACATTTGATGGATGCTCTACRAGCCTCATGTACTACAGCTG 2140
Qy	361 GAAAGAAATATCTACTGCTGTTCATATATATGTTGCTCTCA ----- 403
Db	2141 GAAAGAAATATCTACTGCTGTTCATATATATGTTGCTCTCAAGTGCAGTTCGTAATA 2200
Qy	404 ----- 403
Db	2201 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGG 2260
Qy	404 ----- 403
Db	2261 TTGTAGTATAGACACACAGGGATCTGTGCCAGAGGCTTAAATTTGAAGCGCTTTA 2320
Qy	404 -----GGTCAATAGAAATGACGATCAGGTGGAAGACTCCAAATATCTAGCTT 453
Db	2321 AATATAAATGGGTCAATAGAAATGACGATCAGGTGGAAGACTCCAAATATCTAGCTT 2380
Qy	454 CTCGATATGATTTCAATGACTTAGATCGTGTGGGCATCCACGCGTGGTCTCTATGAGGAT 513
Db	2381 CTCGATATGATTTCAATGACTTAGATCGTGTGGGCATCCACGCGTGGTCTCTATGAGGAT 2440
Qy	514 ACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTTCAGGGTTGCTATTGCTGGGG 573
Db	2441 ACCTCTCCCTGATGGCATTAAATGACAGGTCAGATATCTTTCAGGGTTGCTATTGCTGGGG 2500
Qy	574 CCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAGGAAGCTTATATGGGTACCC 633
Db	

Db	2501 CCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTACCC 2560
Qy	634 CTGACCAAGATGACACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCC 693
Db	2561 CTGACCAAGATGACACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCC 2620
Qy	694 CCTCTGAACCAAAATCGTTTACTCTCTTACATGGTTTCTGTGATGAGAAATGTCCTATTTG 753
Db	2621 CCTCTGAACCAAAATCGTTTACTCTCTTACATGGTTTCTGTGATGAGAAATGTCCTATTTG 2680
Qy	754 CACATACAGTATATTACTAGTATTTTGTAGTGGGCTGGAAGCCATATGATTTACAGA 813
Db	2681 CACATACAGTATATTACTAGTATTTTGTAGTGGGCTGGAAGCCATATGATTTACAGA 2740
Qy	814 TCTATCTCAGGAGAGACACAGCATAAGAGTTTCTGAAATCGGGAGAACATATGAACATGC 873
Db	2741 TCTATCTCAGGAGAGACACAGCATAAGAGTTTCTGAAATCGGGAGAACATATGAACATGC 2800
Qy	874 ATCTTTTGGCACTACCTTTCAAGAAAACCTTGGATCAGCTATTGCTCTTAAAGTGATAT 933
Db	2801 ATCTTTTGGCACTACCTTTCAAGAAAACCTTGGATCAGCTATTGCTCTTAAAGTGATAT 2860
Qy	934 AATTTTGACCTGTGAGAACTCTCTGGTATACACCTGGCTATTTAACCAATGAGAGGTT 993
Db	2861 AATTTTGACCTGTGAGAACTCTCTGGTATACACCTGGCTATTTAACCAATGAGAGGTT 2920
Qy	994 TAATCACAAGAAACACAGAAATGATCATCATTTTGATACCTGCCATGTAACATCTAC 1053
Db	2921 TAATCACAAGAAACACAGAAATGATCATCATTTTGATACCTGCCATGTAACATCTAC 2980
Qy	1054 TCCTGAAAATAAATGTGGTGCCATGCAAGGGGTCTACGGTTTGTGGTAGTAATCTAATACC 1113
Db	2981 TCCTGAAAATAAATGTGGTGCCATGCAAGGGGTCTACGGTTTGTGGTAGTAATCTAATACC 3040
Qy	1114 TTAACCCCATGCTCTCAAAATCAAAATGATACATATTCCTGAGAGACCAGCAATACCATA 1173
Db	3041 TTAACCCCATGCTCTCAAAATCAAAATGATACATATTCCTGAGAGACCAGCAATACCATA 3100
Qy	1174 AGAATTACTAAAAAATAAAAAA 1197
Db	3101 AGAATTACTAAAAAATAAAAAA 3124
RESULT 7	
AX354793	AX354793 3143 bp DNA linear PAT 06-FEB-2002
LOCUS	Sequence 1 from Patent WO0179473.
DEFINITION	AX354793
ACCESSION	AX354793.1 GI:18619526
VERSION	
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1
AUTHORS	Meyers, R.A. and Williamson, M.
TITLE	21953, a human prolyl oligopeptidase family member and uses thereof
JOURNAL	Patent: WO 0179473-A 1 25-OCT-2001.
FEATURES	Millennium Pharmaceuticals, Inc. (US)
source	Location/Qualifiers
	1..3143
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
	229..2877
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAD22990.1"
	/db_xref="GI:18619526"
	/translation="MAAAMEIQGLVETADCEENIESODRPLKPFVVERYSWSQ LKKLLADTRKVGHYMAKAPDPMFKVGRNDPDGPHSDRIYVILAMSGENRNTLFTSEI PKTINRAAVLMLSKPKLLOLFOALDQVGMYSREBELAERXIGTIGASIDYHQSG TELFQAGSGTYHVKDGPQGFQOPLBNLJETSCHTRMDPKUCPADPWIAFIHNS DIWISNIVTFRERLJVAVNELANMEEDASAGUATFVJQEFDRISGYMWCPEAET PSGGKILILYEENDESSEVEI THVTSPLSTRADSFRPKGTANPKVTFKMSIEMI DAGRLIDIVDELIOPRILFEQVEYANAGWTFEGYAWISILLDRSQRLQIVLIS PELPIVEDDVMERGLIESVDSPLIITETIDMINIHDIFHVPQSHBEERIEF IPASECKTGRHLXYKITS LKESKYRSRSGGULPDSDFKPIKEEIAITSGEWELGR HGSNIQVDEVRILVPEGTRDSPLERHLIYVSNVPEVTKLDRGSHSCCSQHCED FRYSKSNKPNCHSLYKLSFDEDDTCKTKEFWATILDSAGPLPYTPPEIFSES TGFTLYGCKYRPHDLQPKQKPYTLVFIYGGQVOLVNNRPGVKVPRLNTLASLOY VVUIDNRGCKYRPHDLQPKQKPYTLVFIYGGQVOLVNNRPGVKVPRLNTLASLOY CGYLSMALMQSLDFRVAIAGAPVILMIFDYDTGYRMYGHPDQNGYLGVSAMQ AERFPEPRLHLLGDFDENVHFATISILLSFLVRAGKPYDLQIYPOERHSIRVPES GEHYELHLHYLOENLGSRIALKVI"
CDS	



TFLEAGSGIYHVKGQPGQFTQOPLRPNLVETSPNIRMDPKLPADPDWIAFIHNS  
DIMSNTVTEERRLTVVHNELEEDARSAGVATFVLOEFDRYSYWCMPKAEIT  
PSGKILRLIYEENDESEVEI IHVTSMPLETRADSFRYPKGTANPKVTPKASEIM  
DAEGRIIDVDEKELIQPELIFEGVEYIARAGWTEGKYAWSILLDRSQRLQIVLIS  
PELPIVEDDVMERQRLIESVPSVTPLIYBEETDIWINIHDIFHPVPPQSHEEIEF  
IFASECTGFRHLVYKITSILKESKYRSGGLPAPSDPKCPIKEEIAITSGMEVLGR  
HGSNIQVDEVRRLVYPEGTKDSPLEHLYVWVYVNPGEVRLTRDGRYSHSCCISOHCD  
PFIISKYNOKNHCVSILYKLSSPEDDPTCKTEFWATILDSAGPLPDVYPELFSRPS  
TTGFTLXGMLYKXHDLOPKKYFTVLFIYGGPQVQVNNRPGKYFRNLNTLASLGYV  
VVVDNDSCHRGKLEFAGKYKMGQIEIDDDQVEGLQYLASKRYDFIDLRVGHWSY  
GGYLSMALMQRSDIFRVAIAGAPVTLWTFYDTGYTIRMGPQNEQYGLVGSVAMQ  
AEKPEPSPNRLLLHGFLENVHFAHTSILLSFLVRAGKPYDLQIYPOERHSRVPES  
GEHVELHLLHVLQENLGRIFAAKVI"

ORIGIN

Query Match 86.9%; Score 1040; DB 6; Length 3143;  
Best Local Similarity 89.1%; Pred. No. 9.7e-267;  
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAA 60  
Db 1796 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 1855  
QY 61 ATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAGTC 120  
Db 1856 ATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAGTC 1915  
QY 121 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCC 180  
Db 1916 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCC 1975  
QY 181 TTTCACAGCTATCAAGTCTTGAAGTACGCCCACTTTGCAAAACAAAGAAATTTTGGGCCA 240  
Db 1976 TTTCACAGCTATCAAGTCTTGAAGTACGCCCACTTTGCAAAACAAAGAAATTTTGGGCCA 2035  
QY 241 CATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAGAAATTTTCTCTTTTG 300  
Db 2036 CATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAGAAATTTTCTCTTTTG 2095  
QY 301 AAAGTACTTACGATTTTACATTTGATGGGATGCTTACAAAGCTCATGATCTACAGCCTG 360  
Db 2096 AAAGTACTTACGATTTTACATTTGATGGGATGCTTACAAAGCTCATGATCTACAGCCTG 2155  
QY 361 GAAAGAAATATCTTACTGTGCTGTTTATATATGTTGGTCTCTCA----- 403  
Db 2156 GAAAGAAATATCTTACTGTGCTGTTTATATATGTTGGTCTCTCAGGTGCGATGTTGGTAATA 2215  
QY 404 ----- 403  
Db 2216 ATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCTACGCTCTTAGGTTATGTGG 2275  
QY 404 ----- 403  
Db 2276 TTGTAGTATAGACACAGGGGATCTCTGTCCAGCGGGCTTAATTTGAAGCGGCCTTAA 2335  
QY 404 -----GGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453  
Db 2336 AATATAAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 2395  
QY 454 CTGATATGATTTTCAATTGACTTATGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAT 513  
Db 2396 CTGATATGATTTTCAATTGACTTATGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAT 2455  
QY 514 ACTCTCCCTGATGCGATTAATGACAGGTCAGATATCTTACAGGTTGCTATGCTGGGG 573  
Db 2456 ACTCTCCCTGATGCGATTAATGACAGGTCAGATATCTTACAGGTTGCTATGCTGGGG 2515  
QY 574 CCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAAAGTTATATGGGTGACC 633  
Db 2516 CCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAAAGTTATATGGGTGACC 2575  
QY 634 CTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCC 693

Db 2576 CTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGTGCCATGCAAGCAGAAAAGTTCC 2635  
QY 694 CCTCTGACCAAAATCGTTTACTCTCTTACATGGTTTCTCTGATCAGAAATGTCCATTTTG 753  
Db 2636 CCTCTGACCAAAATCGTTTACTCTCTTACATGGTTTCTCTGATCAGAAATGTCCATTTTG 2695  
QY 754 CACATACCAAGTATATTAATCTAGTCTGAGGCTCGAAAGCCATATGATTTTACAGA 813  
Db 2696 CACATACCAAGTATATTAATCTAGTCTGAGGCTCGAAAGCCATATGATTTTACAGA 2755  
QY 814 TCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGGAGAACATTTATGAATGC 873  
Db 2756 TCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGGAGAACATTTATGAATGC 2815  
QY 874 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATCTCTCTCTTAAAAGTGATAT 933  
Db 2816 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATCTCTCTCTTAAAAGTGATAT 2875  
QY 934 AATTTTGACCTGTGTAGAACTCTCTGTATACACTGGCTATTTAACCAATGAGGAGGTT 993  
Db 2876 AATTTTGACCTGTGTAGAACTCTCTGTATACACTGGCTATTTAACCAATGAGGAGGTT 2935  
QY 994 TAATCAACAGAAAACACAGAAATTTGATCATCACATTTTGTATACCTGCCATGTAACATCTAC 1053  
Db 2936 TAATCAACAGAAAACACAGAAATTTGATCATCACATTTTGTATACCTGCCATGTAACATCTAC 2995  
QY 1054 TCCTGAAAAATAATGTGGTCCATGACGGGCTCTACGGTTTGTGTAGTATCTTAATACC 1113  
Db 2996 TCCTGAAAAATAATGTGGTCCATGACGGGCTCTACGGTTTGTGTAGTATCTTAATACC 3055  
QY 1114 TTAACCCCACTGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCATTA 1173  
Db 3056 TTAACCCCACTGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCATTA 3115  
QY 1174 AGAATTTACTAAAAAATAAAAAAAAAA 1197  
Db 3116 AGAATTTACTAAAAAATAAAAAAAAAA 3139

RESULT 8  
AR651452  
LOCUS AR651452  
DEFINITION Sequence 2 from patent US 6881564.  
ACCESSION AR651452  
VERSION AR651452.1 GI:62795938  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3120)  
AUTHORS Abbott,C.A. and Gorrell,M.D.  
TITLE Dipeptidyl peptidases  
JOURNAL Patent: US 6881564-A 2 19-APR-2005;  
The University of Sydney; Sydney;  
AUX;

FEATURES  
source Location/Qualifiers  
1..3120  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 86.5%; Score 1036; DB 6; Length 3120;  
Best Local Similarity 89.0%; Pred. No. 1.1e-265;  
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;  
QY 1 ATTTTGAAGGCCAACAAAGACTCCCTTTTATAGAGCATCACCTGTACGTAGTACGTACGTAA 60  
Db 1781 ATTTTGAAGGCCAACAAAGACTCCCTTTTATAGAGCATCACCTGTACGTAGTACGTAA 1840  
QY 61 ATCTTGAGAGGTGACAAAGCTGACGTCGCGGTGCTACTCAATTTCTTGTGCTCATCAGTC 120  
Db 1841 ATCTTGAGAGGTGACAAAGCTGACGTCGCGGTGCTACTCACATTTCTTGTGCTCATCAGTC 1900

121 AGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAAATCCACATGTGTGTCCC 180  
1901 AGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAAATCCACATGTGTGTCCC 1960  
181 TTTTAAAGATCTCAAGTCTGGAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGGCCA 240  
1961 TTTTAAAGATCTCAAGTCTGGAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGGCCA 2020  
241 CCATTTTGGATTCAGCAGAGTCTCTCTCTGAGTATATCTCTCCAGCAAAATTTTCTCTTTTG 300  
2021 CCATTTTGGATTCAGCAGAGTCTCTCTCTGAGTATATCTCTCCAGCAAAATTTTCTCTTTTG 2080  
301 AAAGTACTACTGGAATTTACATGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTG 360  
2081 AAAGTACTACTGGAATTTACATGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTG 2140  
361 GAAAGAAATATCTCTACTGTGCTGCTTCAATATATGGTGGTCTCA----- 403  
2141 GAAAGAAATATCTCTACTGTGCTGCTTCAATATATGGTGGTCTCAAGTGCAGTTCAGTGAATA 2200  
404 ----- 403  
2201 ATCGTTTAAAGAGTCAAGTATTTCCGGCTTGAATACCTAGCCTCTAGGTATATGCG 2260  
404 ----- 403  
2261 TTGTAGTATAGACAACAGGGGATCTGTCAACGAGGGGCTTAATTTGAAGGCGCCTTTA 2320  
404 -----GGGTCAAATAGAAATTTGACGATCAGGTGGAAGGATCCCAATATCTAGCTT 453  
2321 AATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGGATCCCAATATCTAGCTT 2380  
454 CTCGATATGATTTCAATGACTTATAGTCTGTGGGATCCACGGCTGGTCTCATGGAGGAT 513  
2381 CTCGATATGATTTCAATGACTTATAGTCTGTGGGATCCACGGCTGGTCTCATGGAGGAT 2440  
514 ACCTCTCCCTGATGGCAATTAATGACAGGTCAGATATCTTCAGGGTGTCTATTCCTGGGG 573  
2441 ACCTCTCCCTGATGGCAATTAATGACAGGTCAGATATCTTCAGGGTGTCTATTCCTGGGG 2500  
574 CCCAGTCACTCTGTGGATCTTATGATACAGGATACACGGAACGTTATATGGGTCAAC 633  
2501 CCCAGTCACTCTGTGGATCTTATGATACAGGATACACGGAACGTTATATGGGTCAAC 2560  
634 CTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAAAAGTTCC 693  
2561 CTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAAAAGTTCC 2620  
694 CCTCTGAACCAATTCGTTTACTGCTCTTACATGTTTCTCTGATGAGAATGTCATTTTG 753  
2621 CCTCTGAACCAATTCGTTTACTGCTCTTACATGTTTCTCTGATGAGAATGTCATTTTG 2680  
754 CACATACAGTATATTACTGAGTCTTTTGTAGTGGGCTGGAAAGCCATATGATTACAGA 813  
2681 CACATACAGTATATTACTGAGTCTTTTGTAGTGGGCTGGAAAGCCATATGATTACAGA 2740  
814 TCTATCTCAGGAGACACAGATGAAGTTCCTGAATCGGGAGACATTTATGAATGTC 873  
2741 TCTATCTCAGGAGACACAGATGAAGTTCCTGAATCGGGAGACATTTATGAATGTC 2800  
874 ATCTTTTGCATCTCTCAAGAAACCTTTGGATCAAGTATTCGCTCTCAAAAGTATAT 933  
2801 ATCTTTTGCATCTCTCAAGAAACCTTTGGATCAAGTATTCGCTCTCAAAAGTATAT 2860  
934 AATTTTGCATCTCTGATAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTT 993  
2861 AATTTTGCATCTCTGATAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTT 2920  
994 TAATCAACAGAAACACAGAAATTTGATCATCATATTTGATACCTGCCATGTAACATCTAC 1053  
2921 TAATCAACAGAAACACAGAAATTTGATCATCATATTTGATACCTGCCATGTAACATCTAC 2980  
1054 TCCTGAAATAAATGTGGTGCATGAGGGGTCTACGGGTTTGTGGTAGTATCTAATACC 1113

2981 TCCTGAAAAATAATGTGGTGCATGAGGGGTCTAGGTTTGTGGTAGTATCTAATACC 3040  
1114 TTAACCCCAATGCTCTCAAAATCAATGATACATATTTCTGAGAGACCAGCAATACCATA 1173  
3041 TTAACCCCAATGCTCTCAAAATCAATGATACATATTTCTGAGAGACCAGCAATACCATA 3100  
1174 AGAATTACTAAAAA 1193  
3101 AGAATTACTAAAAA 3120  
BC040203 4535 bp mRNA linear PRI 07-OCT-2003  
IMAGE:4824813, complete cds.  
BC040203  
BC040203.1 GI:26007916  
MGC  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 4535)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Spaplen, M., Soares, M.B., Bonaldi, A.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raja, S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Worley, K.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Rodrigues, S.,  
Bouffard, G.C., Blakeley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Bucknerfield, Y.S., Krzywinski, M.I., Skalek, U., Smaluk, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 4535)  
Strausberg, R.  
Direct Submission  
Submitted (27-NOV-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
Toshuyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisegeed, H.,  
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
A.N., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 71 Row: k Column: 23.

FEATURES	Location/Qualifiers		Db	3496	AAAGTACTACTGGATTTTACATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG	3555
	1..4535					
source	/organism="Homo sapiens"		QY	361	GAAAGAAATATCCTACTGTGCTCTTCAATATATGTTGGTGCCTCA	403
	/mol_type="mRNA"					
	/db_xref="taxon:9606"		Db	3556	GAAAGAAATATCCTACTGTGCTCTTCAATATATGTTGGTGCCTCAGGTGCAGTTGGTGAATA	3615
	/clone="MGC:42570 IMAGE:4824813"					
	/tissue_type="Testis"		QY	404	-----	403
	/clone_lib="NIH_MGC_97"					
	/lab_host="DH10B"		Db	3616	ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTTATGTGG	3675
	/note="Vector: pBluescript"					
gene	1..4535		QY	404	-----	403
	/gene="DPP8"					
	/note="synonyms: FLJ20283, FLJ14920, MGC26191"		Db	3676	TTGTAGTATAGACAAACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGCGCCTTTA	3735
	/db_xref="GeneID:54878"					
	/db_xref="MIM:606819"		QY	404	-----GGGTCAAATAGAAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTT	453
	1629..4277					
CDS	/gene="DPP8"		Db	3736	AATATAAAATGGGTCAAATAGAAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTT	3795
	/codon_start=1					
	/product="DPP8 protein"		QY	454	CTCGATATGATTTCAATTGACTTAGATTGCGTGTGGGCATCCACGGCTGGTCTTATGGAGGAT	513
	/protein_id="AAH40203.1"					
	/db_xref="GI:26007917"		Db	3796	CTCGATATGATTTCAATTGACTTAGATTGCGTGTGGGCATCCACGGCTGGTCTTATGGAGGAT	3855
	/db_xref="GeneID:54878"					
	/db_xref="MIM:606819"		QY	514	ACCTCTCCCTGATGGGCATTAATGTCACAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGG	573
	/translations="MAAAMETQLGVETFAEDCEENIESODRPKLPBFFVYRYSWQ					
	LKLLADTRYKHGYMMAKAPHDPMFVKRNDPDGHSRIYYLAWSGENRENTILFYSEI		Db	3856	ACCTCTCCCTGATGGGCATTAATGTCACAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGG	3915
	PKTINRAAAMLSWKLPLDLFOATLDYGRNLSREBELRKRRIKTVGIASVDYHQSFG					
	TFLFOAGSGIYHVKDGPGQFTQOLPRNLYVETSPNTRMDPKLCPADPWIAFIHGN		QY	574	CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC	633
	DIWISNIVREERLYYVNLANMEEDARSAGVATFVLOEFPDRYSYVWCPEAET					
	PSGKILRLIDENDESEVEIIHVTSPLMETRADSPRIKPTGTANPKVTKWSEIMI		Db	3916	CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC	3975
	DAEGRIIDVDEKLIQFFELFGEVYIARAGWTPEGKYAWSILLDRSQRLOIVLIS					
	PELFIPEVDDVMRORLIESVDSVPLIIYBEETDIIWINDIHFVFPQSHBEIEF		QY	634	CTGACCCAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCC	693
	IPASECTGPHLYKITSILKESKYKSSGGLPAPDFKCPKEEIAITSGEWEVLGR					
	HGSNIQVDEVRILYFECKDPSLEHLVYVSNPGEVETLDRGSHSCCISQHCD		Db	3976	CTGACCCAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCC	4035
	FTFISYQNKPHCVSLYKLSPPEDDPTCKTFEAWATILDSAGLPDYTPPELFSFES					
	TGTFTLYMLKPHDLQPGKPYTFLFIYGGPQVLVNNRFGKYFRNLNTLASLGV		QY	694	CCTCTGAACCAAATCGTITTACTGCTCTTCACTGCTTCTTCACTGATGAGAATGTCCATTTTG	753
	VWIDNRGSHQRLKEPKYKMGQIEIDQVEGLQYLASRYDFIDLRVGHGWSY					
	GGVLSLWALMORSDI PRVAIAGAPVTLMIFDTGVTERYMGHPDQNRQGYVLGSAVMO		Db	4036	CCTCTGAACCAAATCGTITTACTGCTCTTCACTGCTTCTTCACTGATGAGAATGTCCATTTTG	4095
	AEKFPSPNRLLLHGFLENVHFAHTSILLSFLVRAGKPYDLQIYQERHSIRVPES					
	GEHYELHLHYLOENIGSLRAALKVI"		QY	754	CACATACCACTATATTAATCTAGTCTGAGTCTTCTTCTGAGGGCTGGAAGCCATATGATTTACAGA	813
	3207..4196					
misc_feature	/gene="DPP8"		Db	4096	CACATACCACTATATTAATCTAGTCTTCTTCTGAGTCTTCTTCTGAGGGCTGGAAGCCATATGATTTACAGA	4155
	/note="DAP2; Region: Dipeptidyl					
	aminopeptidases/acylaminoacyl-peptidases [Amino acid		QY	814	TCTATCTCTCAGGAGAGACACAGCATAAAGAGTTCTTCTGAATCGGGAGAACATTTATGAATGC	873
	transport and metabolism]"					
	/db_xref="CDD:COG1506"		Db	4156	TCTATCTCTCAGGAGAGACACAGCATAAAGAGTTCTTCTGAATCGGGAGAACATTTATGAATGC	4215
ORIGIN			QY	874	ATCTTTTGCACTACTCTTCAAGAAAACCTTGGATCAGTATTGCTGCTCTTAAAAGTGATAT	933
	Query Match					
	Best Local Similarity		Db	4216	ATCTTTTGCACTACTCTTCAAGAAAACCTTGGATCAGTATTGCTGCTCTTAAAAGTGATAT	4275
	Matches 1193; Conservative					
	0; Mismatches		QY	934	AATTTTGACCTGTGTAGAACTCTCTGGTATACATCTGGCTATTCTTAAACCAATGAGAGGTT	993
	1; Gaps					
	1		Db	4276	AATTTTGACCTGTGTAGAACTCTCTGGTATACATCTGGCTATTCTTAAACCAATGAGAGGTT	4335
	ATTTTGAAGGCACAAAGACTCCCTTTAGAGATCACCTGTACGTAGTACGTACGTAA					
	60		QY	994	TAATCAACAGAAAACACAGAAATTTGATCAATTTTTGTATACCTGCCATGTAACTCTAC	1053
	ATTCCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTTCACATTTCTTGTGTCATCAGTC					
	120		Db	4336	TAATCAACAGAAAACACAGAAATTTGATCAATTTTTGTATACCTGCCATGTAACTCTAC	4395
	ATTCCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTTCACATTTCTTGTGTCATCAGTC					
	3315		QY	121	AdCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCC	180
	180					
	3316		Db	4396	TCCTGAAAATTAATGTGGTGCATGACGGGCTACGGGTTTGTGGTGTAGTAACTTAATACC	4455
	AdCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCC					
	3375		QY	1054	TCCTGAAAATTAATGTGGTGCCTGACGGGCTACGGGTTTGTGGTGTAGTAACTTAATACC	1113
	240					
	181		QY	4456	TTTAAACCCACATGCTCAAAATCAAATGATACATTTCTTGAGAGACCCAGCAATACCATA	1173
	3376					
	241		Db	4516	AGAATTACTTAAATAAAAAA 1193	4535
	3495					
	3496		Db	4516	AGAATTACTTAAATAAAAAA 4535	
	3495					
	301		QY	360	AAAGTACTACTGGATTTTACATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG	
	360					



Qy	1	A	TTTTGAAGGCA	CCAAAGAC	TCCCTTTT	AGAGCA	TCA	CCTGT	ACGTAGT	CACTAGT	TACGTAA	60	
Db	1261	A	TTTTGAAGGCA	CCAAAGAC	TCCCTTTT	AGAGCA	TCA	CCTGT	ACGTAGT	CACTAGT	TACGTAA	1320	
Qy	61	A	TCCTGGAGAGT	GACAAGGCT	GACTGAC	CGTGGCT	TACT	CACAT	TTCTTGGCT	GCATCAGTC	120		
Db	1321	A	TCCTGGAGAGT	GACAAGGCT	GACTGAC	CGTGGCT	TACT	CACAT	TTCTTGGCT	GCATCAGTC	1380		
Qy	121	A	GCACCTGTGACT	TTCTTTAT	AGTAAGT	ATAGTA	AAAC	CAGAAAG	AATCCAC	ACTGTGTCTCC	180		
Db	1381	A	GCACCTGTGACT	TTCTTTAT	AGTAAGT	ATAGTA	AAAC	CAGAAAG	AATCCAC	ACTGTGTGTCTCC	1440		
Qy	181	T	TTTCAAAAGCT	ATCAAGT	CCTGAAG	ATGAC	CCAACT	TTGCAAAA	CAAAAGG	AAATTTTGGGCCA	240		
Db	1441	T	TTTCAAAAGCT	ATCAAGT	CCTGAAG	ATGAC	CCAACT	TTGCAAAA	CAAAAGG	AAATTTTGGGCCA	1500		
Qy	241	C	CAATTTTGGAT	TACAGGTC	CCTCTT	CCTGACT	TATACT	CCTCC	AGAAATTTTCT	CTTTTG	300		
Db	1501	C	CAATTTTGGAT	TACAGGTC	CCTCTT	CCTGACT	TATACT	CCTCC	AGAAATTTTCT	CTTTTG	1560		
Qy	301	A	AAGTACTACT	CTGGATTT	ACATTTG	TATGGG	ATGCTT	ACAAG	CCTCAT	GATCTACAGCTG	360		
Db	1561	A	AAGTACTACT	CTGGATTT	ACATTTG	TATGGG	ATGCTT	ACAAG	CCTCAT	GATCTACAGCTG	1620		
Qy	361	G	AAAGAAATAT	CCCTACT	GTGTCT	GTTCAT	TATAT	TGGTGGT	CTC	-----	402		
Db	1621	G	AAAGAAATAT	CCCTACT	GTGTCT	GTTCAT	TATAT	TGGTGGT	CTC	-----	1680		
Qy	403	-----	-----	-----	-----	-----	-----	-----	-----	-----	402		
Db	1681	T	ATATCGGTTT	AAAGGAGT	CAAGTAT	TTCCGCTT	GAAAT	ACCT	TAGCCT	CTCTAGGTAT	1740		
Qy	403	-----	-----	-----	-----	-----	-----	-----	-----	-----	402		
Db	1741	G	GTTTGTAGT	GATAGACA	AGGGAT	CCTGT	CTCAC	CGG	GGCTTAA	TTTGAAGGCGCCT	1800		
Qy	403	-----	-----	-----	-----	-----	-----	-----	-----	-----	451		
Db	1801	T	AAATAT	AAAAAT	GGGTCAA	ATAGAAA	TTGAC	GCAT	CAGGT	GGAAGGACT	CCAATATCT	1860	
Qy	452	T	TTCTCGAT	TATGATTT	CATTGACT	TTAGAT	CGTGTGG	GCATC	CA	CGGCTGGT	CCTATGAGG	511	
Db	1861	T	TTCTCGAT	TATGATTT	CATTGACT	TTAGAT	CGTGTGG	GCATC	CA	CGGCTGGT	CCTATGAGG	1920	
Qy	512	A	TACTCT	CCCTGAT	GGCATTAA	TGC	AGAGST	CAGATAT	CTT	CAGGTTGCT	TATTGCTGG	571	
Db	1921	A	TACTCT	CCCTGAT	GGCATTAA	TGC	AGAGST	CAGATAT	CTT	CAGGTTGCT	TATTGCTGG	1980	
Qy	572	G	GGCC	CAGTCACT	CTGTGG	ATCTT	ATAT	GATAC	AGGATAC	CAGGAAC	GTTATATGGGTCA	631	
Db	1981	G	GGCC	CAGTCACT	CTGTGG	ATCTT	ATAT	GATAC	AGGATAC	CAGGAAC	GTTATATGGGTCA	2040	
Qy	632	C	CTGAC	CCAGAA	TGAAC	GGGCTAT	TACT	TAGGAT	CTGTGG	CATC	CAAGCAGAAAAAT	691	
Db	2041	C	CTGAC	CCAGAA	TGAAC	GGGCTAT	TACT	TAGGAT	CTGTGG	CATC	CAAGCAGAAAAAT	2100	
Qy	692	C	CCCTCT	GAA	CCAAAT	CGTTT	ACTGCT	CTTAC	ATGTTT	CCTGG	ATGAAATGTCATTT	751	
Db	2101	C	CCCTCT	GAA	CCAAAT	CGTTT	ACTGCT	CTTAC	ATGTTT	CCTGG	ATGAAATGTCATTT	2160	
Qy	752	T	GCA	TAC	AC	CGATAT	TATTA	CTGAG	TTTTTT	TAGTGAG	GGGCTGGAAG	811	
Db	2161	T	GCA	TAC	AC	CGATAT	TATTA	CTGAG	TTTTTT	TAGTGAG	GGGCTGGAAG	2220	
Qy	812	G	ATCTAT	CTCT	CAGGAG	ACAC	AGATTA	AGGTTT	CTCT	GAATCGG	GAGACATATG	871	
Db	2221	G	ATCTAT	CTCT	CAGGAG	ACAC	AGATTA	AGGTTT	CTCT	GAATCGG	GAGACATATG	2280	
Qy	872	G	CACT	TTTT	TG	CACT	ACCTTT	CAAGAAAA	CC	TTGGAT	CA	CGTATTTGCT	931
Db	2281	G	CACT	TTTT	TG	CACT	ACCTTT	CAAGAAAA	CC	TTGGAT	CA	CGTATTTGCT	2340
Qy	932	A	TAAATTT	TTG	CACT	GTGT	AGAACT	CTCT	GGTAT	TAC	CTGGCTATTT	991	

[illegible]





```
QY 632 CCCTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTT 691
|
|
|
Db 2561 CCCTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTT 2620
|
|
|
QY 692 CCCCTCGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGGATGAGAATGTCATTT 751
|
|
|
Db 2621 CCCCTCGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGGATGAGAATGTCATTT 2680
|
|
|
QY 752 TGCATACACAGATATATCTAGTATTTTGTAGGGCTGGAAGCCATATGATTACA 811
|
|
|
Db 2681 TGCATACACAGATATATCTAGTATTTTGTAGGGCTGGAAGCCATATGATTACA 2740
|
|
|
QY 812 GATCTATCCTCAGGAGAGACACAGCAATAGAGTTCTGAATCGGAGAACATTTAGAACT 871
|
|
|
Db 2741 GATCTATCCTCAGGAGAGACACAGCAATAGAGTTCTGAATCGGAGAACATTTAGAACT 2800
|
|
|
QY 872 GCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTAATGCTGTCTTAAAGTGAT 931
|
|
|
Db 2801 GCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTAATGCTGTCTTAAAGTGAT 2860
|
|
|
QY 932 ATAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACAAATGAGGAG 991
|
|
|
Db 2861 ATAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACAAATGAGGAG 2920
|
|
|
QY 992 TTTAATCAACAGAAAACACAGAAATGATCATACATTTTGATACCTGCCATGTACATCT 1051
|
|
|
Db 2921 TTTAATCAACAGAAAACACAGAAATGATCATACATTTTGATACCTGCCATGTACATCT 2980
|
|
|
QY 1052 ACTCTGAAATTAATGTTGTCATGCGAGGGCTCTACGGTTTGTAGTAATCTAATA 1111
|
|
|
Db 2981 ACTCTGAAATTAATGTTGTCATGCGAGGGCTCTACGGTTTGTAGTAATCTAATA 3040
|
|
|
QY 1112 CTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTCAGAGAGCCAGCAATACCA 1171
|
|
|
Db 3041 CTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTCAGAGAGCCAGCAATACCA 3100
|
|
|
QY 1172 TAAGAAATTAATAAAAAAAAAAAAAA 1197
|
|
|
Db 3101 TAAGAAATTAATAAAAAAAAAAAAAA 3126
|
|
|
```

```
RESULT 14
AX342633 LOCUS AX342633 3106 bp DNA linear PAT 12-JAN-2002
DEFINITION Sequence 30 from Patent W00198468.
ACCESSION AX342633
VERSION AX342633.1 GI:18152030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Elliott,V.S., Gandhi,A.R., Lal,P., Au-Young,J.,
Tribouley,C.M., Delemane,A.M., Baughn,M.R., Nguyen,D.B., Lee,E.A.,
Hafalia,A., Khan,F.A., Walia,N.K., Yao,M.G., Lu,D.A., Patterson,C.,
Tang,Y.T., Walsh,R.T., Azimzai,Y., Ramkumar,J., Xu,Y. and Reddy,R.
JOURNAL Patent: WO 0198468-A 30 27-DEC-2001;
Incite Genomics, Inc. (US)
FEATURES
source
1..3106
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incite ID No: 7160544CB1"
```

## ORIGIN

```
Query Match 86.3%; Score 1033; DB 6; Length 3106;
Best Local Similarity 89.0%; Pred. No. 7.2e-265;
Matches 1190; Conservative 0; Mismatches 0; Indels 147; Gaps 1;
QY 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACCTAGTACGTACGTAA 60
```

```
Db 1770 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAA 1829
|
|
|
QY 61 ATCTCTGGAGAGGTGACAAAGGCTGACCTGACCGTGGCTACTCACATTTCTGCTGCATCAGTC 120
|
|
|
Db 1830 ATCTCTGGAGAGGTGACAAAGGCTGACCTGACCGTGGCTACTCACATTTCTGCTGCATCAGTC 1889
|
|
|
QY 121 AGCACTGTGACTCTCTTTATAAGTAAGTATAGTAAACAGAGAAGATCCACACTGTGTGTCC 180
|
|
|
Db 1890 AGCACTGTGACTCTCTTTATAAGTAAGTATAGTAAACAGAGAAGATCCACACTGTGTGTCC 1949
|
|
|
QY 181 TTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGGCAAAACAAAGGAATTTTGGGCCA 240
|
|
|
Db 1950 TTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGGCAAAACAAAGGAATTTTGGGCCA 2009
|
|
|
QY 241 CCAATTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTG 300
|
|
|
Db 2010 CCAATTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTG 2069
|
|
|
QY 301 AAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGCCCTCATGATCTACAGCCTG 360
|
|
|
Db 2070 AAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGCCCTCATGATCTACAGCCTG 2129
|
|
|
QY 361 GAAAGAAATATCCCTACTGCTGTTTATATATGCTGCTCTCA----- 403
|
|
|
Db 2130 GAAAGAAATATCCCTACTGCTGTTTATATATGCTGCTCTCAGGTGCAAGTTGGTGAATA 2189
|
|
|
QY 404 ----- 403
|
|
|
Db 2190 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTATGTGG 2249
|
|
|
QY 404 ----- 403
|
|
|
Db 2250 TTGTAGTAGTAGACCAACAGGGGATCTGTACCGAGGGCTTAAATTTGAAGGCGCTTTA 2309
|
|
|
QY 404 -----GGGTCAAATAGAAATTGAGGATCAGGTGGAAGGACTCCAATATCTAGCTT 453
|
|
|
Db 2310 AATATAAAATGGGTCAAATAGAAATTGAGGATCAGGTGGAAGGACTCCAATATCTAGCTT 2369
|
|
|
QY 454 CTCGATATGATTTCAATTGACTTTAGATCGTGGGCAATCCACGGCTGCTCTATGAGGAT 513
|
|
|
Db 2370 CTCGATATGATTTCAATTGACTTTAGATCGTGGGCAATCCACGGCTGCTCTATGAGGAT 2429
|
|
|
QY 514 ACCTCTCCCTGATGGCAATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATGTGGGG 573
|
|
|
Db 2430 ACCTCTCCCTGATGGCAATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATGTGGGG 2489
|
|
|
QY 574 CCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 633
|
|
|
Db 2490 CCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 2549
|
|
|
QY 634 CTGACCAAGATGAACAGGGCTATTACTTGGATCTGTGGCCATGCAAGCAGAAAAGTTTCC 693
|
|
|
Db 2550 CTGACCAAGATGAACAGGGCTATTACTTGGATCTGTGGCCATGCAAGCAGAAAAGTTTCC 2609
|
|
|
QY 694 CCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGATGAGAATGTCATTTTG 753
|
|
|
Db 2610 CCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGATGAGAATGTCATTTTG 2669
|
|
|
QY 754 CACATACAGTATATTAATGAGTTTTTGTAGGGCTGGAAGCCATATGATTTTACAGA 813
|
|
|
Db 2670 CACATACAGTATATTAATGAGTTTTTGTAGGGCTGGAAGCCATATGATTTTACAGA 2729
|
|
|
QY 814 TCTATCTCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAGCAATATTAAGCTGC 873
|
|
|
Db 2730 TCTATCTCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAGCAATATTAAGCTGC 2789
|
|
|
QY 874 ATCTTTTGCACCTACCTTCAAGAAAAACCTTGGATCAGTATTCGCTCTTAAAGTGATAT 933
|
|
|
Db 2790 ATCTTTTGCACCTACCTTCAAGAAAAACCTTGGATCAGTATTCGCTCTTAAAGTGATAT 2849
|
|
|
QY 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAAAATGAGGAGTT 993
|
|
|
```



Db 2850 AATTTGACCTGTAGAACTCTCTGGTATACACTGGCTATTATTAACCAATGAGAGGTT 2909

QY 994 TAATCAACAGAAACACAGAAATTGATCATCATCATTTTGTATCTCCGATGAACATCTAC 1053

Db 2910 TAATCAACAGAAACACAGAAATTGATCATCATCATTTTGTATCTCCGATGAACATCTAC 2969

QY 1054 TCCTGAAATAAATGTGGTCCCATGACGAGGGTCTACGGTTTGTGTAGTAACTAAATACC 1113

Db 2970 TCCTGAAATAAATGTGGTCCCATGACGAGGGTCTACGGTTTGTGTAGTAACTAAATACC 3029

QY 1114 TTAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGAGAGACCCAGCAATACCATA 1173

Db 3030 TTAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGAGAGACCCAGCAATACCATA 3089

QY 1174 AGAATTACTTAAAAAAA 1190

Db 3090 AGAATTACTTAAAAAAA 3106

RESULT 15

BC030688

LOCUS BC030688 3125 bp mRNA linear PRI 28-JUL-2005

DEFINITION Homo sapiens dipeptidylpeptidase 8, transcript variant 4, mRNA

ACCESSION (cDNA clone MGC:26191 IMAGE:4822550), complete cds.

VERSION BC030688.2 GI:34190028

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Berge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenner, C.M., Schuler, G.D., Altschuler, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F., Diatchenko, L., Marudina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Casavant, T.L., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Woxley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakesley, R.W., Young, A.C., Shevchenko, Y., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

CONSRTH

TITLE Mammalian Gene Collection Program Team

PUBMED Generation and initial analysis of more than 15,000 full-length

AUTHORS human and mouse cDNA sequences

QY 12477932 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 3125)

NIH MGC Project

Direct Submission

Submitted (24-MAY-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 25, 2003 this sequence version replaced gi:21265132.

Contact: MGC help desk

Email: [cgabbe-x@mail.nih.gov](mailto:cgabbe-x@mail.nih.gov)

Tissue Procurement: Miklos Pavlovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

cDNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: [amadansystemsbio.org](mailto:amadansystemsbio.org)

Anup Madan, Jessica Fahey, Brin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 33 Row: d Column: 5

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 37577092. Location/Qualifiers

1. 3125

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:26191 IMAGE:4822550"

/tissue\_type="Testis"

/clone\_lib="NIH MGC 97"

/lab\_host="DH10B"

/note="Vector: pBluescriptR"

1. 3125

/gene="DPP8"

/notes="synonyms: FLJ20283, FLJ14920, MGC26191, MSTP141"

/db\_xref="GeneID:54878"

/db\_xref="MIM:606819"

319\_-2862

/gene="DPP8"

/codon\_start=1

/product="dipeptidyl peptidase 8, isoform 4"

/protein\_id="AAH30688.3"

/db\_xref="GI:54114905"

/db\_xref="GeneID:54878"

/db\_xref="MIM:606819"

/translations="MWKSEQWKIKSGKCNMAAMETQLGVETFDACBENTESQD RPKLEFYVERYSQKLKADATRYHGMKAPDFMFVKNYDMDGDPGSHDRYYL AMGENRNFYSIEPKTINRAAALMSWKLDFDQTDYKNDYSGPDSRLRKR IGTVGASDYDHQSGTFLPQGGSGYHVKGQPGFTQQLPRLNVLVETSCPNRMOP KLCPADYDFATHSNDIMSNIVTREERLTYVHNEANMEEDARSAGVATFVQEE FDRYSGWCMCPKAETTPSGGKILRLDYBENDESEVEIIHVTSPLMETRADSPRYPK GTANPKVTQKMSIMDASGRILVDIKLOPFELFGEVETARAGTEPGKYAMS ILLDRSQRLQIVLISPELFIPEVDVMEQRQLIESVPSVTSVLIIEETDWINIH DIFHFPQSEHEIEFISECKTGRHLKITSILKSKYSSGGLPAPSDFKPIL KESATISGEMVQLHGHSNIQVDEVRLLVPEGKDSPLSEHLHYVVSYNVDFKRL DEATYSHSCISQCHDFISKYSNQKNPHCVSLYKLSPEDDPTCKKEFWATLDSA GLPFDYPTPEIFESTTGTFTGLYKPHDLQPGKGYPTVLYFGGQVLVNNRFK GVYFRNLTLASLGYVVVVDNRGSHRGLKPEGAFKYMVAIAGAPVTLWIFDVTGY TERYMGHPDQNEQGYLGSVAMQAEKFFSEPNRLLHLHFLDENVHAFTSILLSFLV RAGKPYDLQIYQPERHSIRVPESGSEHYELHLLHYLQENLGSRIAALKVI"

ORIGIN

Query Match 84.9%; Score 1016.2; DB 8; Length 3125;

Best Local Similarity 91.7%; Pred. No. 2.2e-260;

Matches 1098; Conservative 0; Mismatches 93; Indels 6; Gaps 2;

QY 1 ATTTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTACGTACGTAA 60

Db 1934 ATTTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTACGTAA 1993

QY 61 ATCTGAGAGGTGACAGAGGTGACTGACCGTGGCTACTCACATCTTGTGTCATCAGTC 120

Db 1994 ATCTGAGAGGTGACAGAGGTGACTGACCGTGGCTACTCACATCTTGTGTCATCAGTC 2053

QY 121 AGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTC 180

Db 2054 AGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTC 2113

QY 181 TTTTAAAGCTATCAAGTCTCGAAGATGACCCAACTTGGCAAAACAAAGAAATTTGGGGCA 240

Db 2114 TTTTAAAGCTATCAAGTCTCGAAGATGACCCAACTTGGCAAAACAAAGAAATTTGGGGCA 2173

QY 241 CCATTTTGGATTGACGAGGTCTCTTCTCTGACTATATCTCTCCACGAAATTTTCTTTTG 300

Db 2174 CCATTTTGGATTGACGAGGTCTCTTCTCTGACTATATCTCTCCACGAAATTTTCTTTTG 2233

QY 301 AAGTACTACTGGATTATCAATTGATGGAGTGTCTCTACAAGCCTCATGATCTACAGCCTG 360

```

Db      2234 AAAGTACTACTGGATTTACATGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTG 2293
QY      361 GAAAGAAATATCTTACTGTCTGTGTTCAATATATGGTGGTCTCAGGGTCAAAATAGAAATTG 420
Db      2294 GAAAGAAATATCTTACTGTCTGTGTTCAATATATGGTGGTCTCAGGGTCAAGTTGG--TGA 2350
QY      421 ACGATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATC 480
Db      2351 AATAACGGTTTAAAGGAGTCAAGTATTTCCGCT---TGAATACCCCTAGCCTCTCTAGGTT 2407
QY      481 GTGTGGGCATCCACGGCTGTCTCTATGAGGAGTACTCTCTCCTGATGGCAATTAATGCAGA 540
Db      2408 ATGTGGTTGTAGTGTATAGACACAGGGGATCTCTGCACGGGGCTTAATTTGAAGGCG 2467
QY      541 GGTCAAGATATCTTTCAAGGTTGTCTATGTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 600
Db      2468 CCTTTAAATATAAATGGTTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 2527
QY      601 ATACAGGATACACGGAAACGTTATATGGGTACCCCTGACAGAAATGAACAGGGCTTATTA 660
Db      2528 ATACAGGATACACGGAAACGTTATATGGGTACCCCTGACAGAAATGAACAGGGCTTATTA 2587
QY      661 TAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720
Db      2588 TAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 2647
QY      721 TACATGGTTTCTCGATGAGATGCCATTTTGGACATACCAGTATATTAATGAGTTT 780
Db      2648 TACATGGTTTCTCGATGAGATGCCATTTTGGACATACCAGTATATTAATGAGTTT 2707
QY      781 TAGTGGGGCTGGAAGGCCATATGATTTACAGATCTATCTCCTCAGGAGACACAGCATAA 840
Db      2708 TAGTGGGGCTGGAAGGCCATATGATTTACAGATCTATCTCCTCAGGAGACACAGCATAA 2767
QY      841 GAGTTCCTGTAATCGGAGAACATTAATGAACCTGCTTTTGGACCTACCTTCAAGAAAACC 900
Db      2768 GAGTTCCTGTAATCGGAGAACATTAATGAACCTGCTTTTGGACCTACCTTCAAGAAAACC 2827
QY      901 TTGGATCACGTATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG 960
Db      2828 TTGGATCACGTATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG 2887
QY      961 TATACACTGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACACAGAAATGATC 1020
Db      2888 TATACACTGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACACAGAAATGATC 2947
QY      1021 ATCACATTTTGATACCTGCCATGTAACTACTCTGAAATAAATGTTGGTGGCCATGCA 1080
Db      2948 ATCACATTTTGATACCTGCCATGTAACTACTCTGAAATAAATGTTGGTGGCCATGCA 3007
QY      1081 GGGGTCTACGGTTTGTGTAGTAACTCTAATACCTTAACCCACATGCTCAAAATCAAAATG 1140
Db      3008 GGGGTCTACGGTTTGTGTAGTAACTCTAATACCTTAACCCACATGCTCAAAATCAAAATG 3067
QY      1141 ATACATATTTCTCAGAGACCCAGCAATACCAATGAAATTTACTTAAAAA 1197
Db      3068 ATACATATTTCTCAGAGACCCAGCAATACCAATGAAATTTACTTAAAAA 3124

```

Search completed: May 4, 2006, 18:52:10  
Job time : 5977.21 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: May 4, 2006, 02:50:44 ; Search time 674.127 Seconds  
(without alignments)  
11834.025 Million cell updates/sec  
Title: US-10-825-632-4  
Perfect score: 1197  
Sequence: 1 atttgaaggccacaaagac.....ttactaataaaacaaaaa 1197  
Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4996997 seqs, 332346308 residues  
Total number of hits satisfying chosen parameters: 9993994  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	1197	100.0	1197 4 AAC85695 Nucleotid
2	1179	98.5	4676 6 ABK83331
3	1040	86.9	2083 7 ADL1222 Human CD-
4	1040	86.9	2797 12 ADL13374 Human ste
5	1040	86.9	3143 6 AAH99934
6	1036	86.5	3120 4 AAC85694
7	1036	86.5	3120 6 AAD38956
8	1033.2	86.3	4309 6 ABK83328
9	1033.2	86.3	4829 6 ABK83327
10	1033	86.3	3106 6 ABK12892
11	1030	86.0	2929 10 ACA92425
12	1021	85.3	1821 6 ABV76411
13	1013.6	84.7	4685 6 ABK83332
14	1005.6	84.0	2952 10 ACA92421
15	999.6	83.5	2161 4 AAH15009
16	897.6	75.0	2251 10 ADE79035
17	881	73.6	1669 4 AAC85696
18	879	73.4	2668 6 ABN59775
19	879	73.4	2842 6 ABN59774

20	873	72.9	4523	6	ABK83325
21	863.6	72.1	2463	10	ACA92424
22	863.6	72.1	2510	6	AAD23843
23	840.8	70.2	2702	13	ADT04072
24	791.4	66.1	873	4	AAF81719
25	791.4	66.1	925	6	ABL90148
26	778	65.0	2649	14	ADV43981
27	776.4	64.9	2649	8	ABX12255
28	776.4	64.9	2671	6	ABK83322
29	775	64.7	2643	6	AAH99935
30	633.4	52.9	823	6	ABK30401
31	619	51.7	2349	14	ADV43982
32	523.6	43.7	587	4	AAH12830
33	508	42.4	2696	12	ADI16386
34	459	38.3	1083	4	AAC85697
35	380.4	31.8	561	4	AAJ00876
36	326.8	27.3	612	6	ABN86324
37	308.8	25.8	502	9	ACH32373
38	237.2	19.8	3287	6	AAD38955
39	214.4	17.9	2660	6	ADI16689
40	214.4	17.9	2660	6	ADI16687
41	212.8	17.8	662	4	AAK92083
42	212.8	17.8	662	4	AAK93366
43	212.8	17.8	662	12	ADI29793
44	212.8	17.8	662	12	ADI28510
45	212.8	17.8	1048	4	AAS41004

ALIGNMENTS

RESULT 1  
AAC85695  
ID AAC85695 standard; cDNA; 1197 BP.  
XX  
AC AAC85695;  
XX  
DT 29-JUN-2001 (first entry)  
XX  
DE Nucleotide sequence encoding human DPP8 524Phe-657Gln+707Gly-882Ile.  
XX  
KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;  
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;  
KW growth hormone deficiency; glucose level; mucosal regeneration;  
KW non-insulin dependent diabetes mellitus; glucose intolerance;  
KW immunosuppression; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 3..935  
FT /\*tag= a  
FT /product= "Human DPP8 524Phe-657Gln+707Gly-882Ile"  
XX  
PN WO200119866-A1.  
XX  
PD 22-MAR-2001.  
XX  
PF 11-SEP-2000; 2000WO-AU001085.  
XX  
PR 10-SEP-1999; 99AU-00002762.  
PR 18-FEB-2000; 2000AU-00005709.  
XX  
PA (UNSY ) UNIV SYDNEY.  
XX  
PI Abbott CA, Gorell MD;  
XX  
DR WPI; 2001-281520/29.  
XX  
DR P-PSDB; AAB47188.  
XX  
PT New human dipeptidyl aminopeptidase (DPP8) useful for cleaving  
PT substrates, identifying inhibitors of DPP8 catalytic activity which have  
PT therapeutic uses, and for detecting activated T cells.

Claim 21; Page 69-70; 78pp; English.

The sequences given in AAC85695-97 encode fragments of human dipeptidyl aminopeptidase (DPP8). DPP8 has substrate specificity for H-Ala-Pro-PNA, H-Gly-Pro-PNA and H-Arg-Pro-PNA. Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase, because it is capable of hydrolyzing the peptide bond C-terminal to proline in each of these compounds. DPP8 is homologous with human DPP4. DPP8 is useful for cleaving a substrate, and for detecting an activated T cell which involves measuring the level of DPP8 gene expression in a T cell. The level of DPP8 expression is detected by detecting the amount of DPP8 RNA in the cell. It is also useful for identifying a molecule capable of inhibiting the cleavage of the substrate by DPP8. Molecules identified as inhibiting DPP8 catalytic activity may be useful for treating diarrhoea, growth hormone deficiency, lowering glucose levels in non-insulin dependent diabetes mellitus and other disorders involving glucose intolerance, enhancing mucosal regeneration and as immunosuppressants

Sequence 1197 BP; 357 A; 258 C; 242 G; 340 T; 0 U; 0 Other;

Query Match 100.0%; Score 1197; DB 4; Length 1197;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTTGAAGGCGCAACAAAGACTCCCTTTAGAGCATCAGCTGACGATGAGTACGTTAGCTAA 60  
 DB 1 ATTTGAAGGCGCAACAAAGACTCCCTTTAGAGCATCAGCTGACGATGAGTACGTTAGCTAA 60  
 QY 61 ATCTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCATTCTTGTGTCATCAGTC 120  
 DB 61 ATCTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCATTCTTGTGTCATCAGTC 120  
 QY 121 AGCACTGTGACTCTTTTATAAGTAAAGTATAGTAAACCAAGAAATCCACATGTCCTCC 180  
 DB 121 AGCACTGTGACTCTTTTATAAGTAAAGTATAGTAAACCAAGAAATCCACATGTCCTCC 180  
 QY 181 TTTACAGCTATCAAGTCTGAGATGACCCAACTTGCNAAACAAAGAAATTTTGGCCA 240  
 DB 181 TTTACAGCTATCAAGTCTGAGATGACCCAACTTGCNAAACAAAGAAATTTTGGCCA 240  
 QY 241 CCATTTGGATTCAGAGGCTCTCTCTGACTATATCTCCAGAAATTTCTCTTTG 300  
 DB 241 CCATTTGGATTCAGAGGCTCTCTCTGACTATATCTCCAGAAATTTCTCTTTG 300  
 QY 301 AAAGTACTGATTTACATTTGATGGATGCTCTACAGCTCATGATCTACAGCTG 360  
 DB 301 AAAGTACTGATTTACATTTGATGGATGCTCTACAGCTCATGATCTACAGCTG 360  
 QY 361 GAAAGAAATATCTACTGCTGTGTTTCATATATGTTGTTCTCAGGGTCAAAATAGAAATTG 420  
 DB 361 GAAAGAAATATCTACTGCTGTGTTTCATATATGTTGTTCTCAGGGTCAAAATAGAAATTG 420  
 QY 421 ACGATCAGGTGGAAGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATC 480  
 DB 421 ACGATCAGGTGGAAGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATC 480  
 QY 481 GTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAATAGCAGA 540  
 DB 481 GTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAATAGCAGA 540  
 QY 541 GGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATG 600  
 DB 541 GGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATG 600  
 QY 601 ATACAGGATACACGGAACTTATATGAGGTACCTGACCAAGAAATGAACAGGGCTATTACT 660  
 DB 601 ATACAGGATACACGGAACTTATATGAGGTACCTGACCAAGAAATGAACAGGGCTATTACT 660  
 QY 661 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720  
 DB 661 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720

QY 721 TACATGGTTTCTGGATGAGATGTCATTTTGCACATACACAGTATATATTCTAGTGTTTT 780  
 DB 721 TACATGGTTTCTGGATGAGATGTCATTTTGCACATACACAGTATATATTCTAGTGTTTT 780  
 QY 781 TAGTGAGGCTGGAAAGCCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAA 840  
 DB 781 TAGTGAGGCTGGAAAGCCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAA 840  
 QY 841 GAGTTCCTGAATCGGAGAACCATATATGAATCTGATCTTTTGCACCTTCAAGAAAAACC 900  
 DB 841 GAGTTCCTGAATCGGAGAACCATATATGAATCTGATCTTTTGCACCTTCAAGAAAAACC 900  
 QY 901 TTGATCACATATTTGCTGCTTAAAGTATATAATTTTGAACCTGTAGAACTCTCTGG 960  
 DB 901 TTGATCACATATTTGCTGCTTAAAGTATATAATTTTGAACCTGTAGAACTCTCTGG 960  
 QY 961 TATACACTGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATTGATC 1020  
 DB 961 TATACACTGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATTGATC 1020  
 QY 1021 ATCACATTTTGTATACCTGCTGATCACTACTCTGAAATATAATGTTGTCATGCA 1080  
 DB 1021 ATCACATTTTGTATACCTGCTGATCACTACTCTGAAATATAATGTTGTCATGCA 1080  
 QY 1081 GGGTCTACCGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATG 1140  
 DB 1081 GGGTCTACCGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATG 1140  
 QY 1141 ATCATATTTCTCAGAGAGCCAGGATACCATTAAGATTAATTAATAAAAAA 1197  
 DB 1141 ATCATATTTCTCAGAGAGCCAGGATACCATTAAGATTAATTAATAAAAAA 1197

RESULT 2  
 ABK83331  
 ID ABK83331 standard; cDNA; 4676 BP.  
 AC ABK83331;  
 DT 12-AUG-2002 (first entry)  
 XX cDNA encoding human DPP8-1 splice variant #7.  
 DE Human; serine protease; dipeptidyl peptidase IV-related protein; DPP8;  
 KW DPP8; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;  
 KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;  
 KW heart failure; hypertension; urinary retention; osteoporosis; cancer;  
 KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;  
 KW dyskinesia; reproductive disorder; inflammatory disorder;  
 KW metabolic disorder; gene; ss.  
 XX Homo sapiens.  
 XX WO200231134-A2.  
 XX 18-APR-2002.  
 XX 12-OCT-2001; 2001WO-US031874.  
 XX 12-OCT-2000; 2000US-0240117P.  
 XX (PERR ) FERRING BV.  
 XX Qi S, Akinsanya KO, Riviere PJ, Junien J;  
 WPI: 2002-444178/47.  
 P-PSDB; ABG61600.  
 XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding  
 the proteins useful for treating e.g. fungal, bacterial, protozoan and  
 viral infections, cancers, allergies, neurological disorders, or pain.  
 XX Disclosure; Page 72-73; 113pp; English.

XX The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins  
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic  
CC acids encoding them are useful for treating infections such as fungal,  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABR83322  
CC -ABK83343 encode human DPRP proteins  
XX  
SQ Sequence 4676 BP; 1424 A; 859 C; 979 G; 1414 T; 0 U; 0 Other;  
  
Query Match 98.5%; Score 1179; DB 6; Length 4676;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1193; Conservative 0; Mismatches 0; Indels 4; Gaps 1;  
  
Qy 1 ATTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAA 60  
Db 1781 ATTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAA 1840  
  
Qy 61 ATCTCGAGAGGTGACAGGCTGACTGACCGTGGTACTCACATTTCTGCTGCATCAGTC 120  
Db 1841 ATCTCGAGAGGTGACAGGCTGACTGACCGTGGTACTCACATTTCTGCTGCATCAGTC 1900  
  
Qy 121 AGCACTGTGACTTCTTTAAGTAAGTATAGTACCAAGAGAATCCACACTGTGTGCC 180  
Db 1901 AGCACTGTGACTTCTTTAAGTAAGTATAGTACCAAGAGAATCCACACTGTGTGCC 1960  
  
Qy 181 TTTTACAAGCTATCAAGTCTCTGAAGATGCCCAACTTGCACCAAGAGAAATTTTGGGCCA 240  
Db 1961 TTTTACAAGCTATCAAGTCTCTGAAGATGCCCAACTTGCACCAAGAGAAATTTTGGGCCA 2020  
  
Qy 241 CCAATTTTGGATTCAGAGGCTCTCTTCTGATATATCTCTCCGAGAAATTTCTCTTTTG 300  
Db 2021 CCAATTTTGGATTCAGAGGCTCTCTTCTGATATATCTCTCCGAGAAATTTCTCTTTTG 2080  
  
Qy 301 AAGTACTACTGATTTACATTTATGGATGCTCTACAGCTCATGATCAGCGTG 360  
Db 2081 AAGTACTACTGATTTACATTTATGGATGCTCTACAGCTCATGATCAGCGTG 2140  
  
Qy 361 GAAAGAAATATCTACTGCTGCTTCAATATATGGTGGTCCCTCAGGGTCAATAGAAATG 420  
Db 2141 GAAAGAAATATCTACTGCTGCTTCAATATATGGTGGTCCCTCAGGGTCAATAGAAATG 2196  
  
Qy 421 AGCATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGACTTAGTC 480  
Db 2197 AGCATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGACTTAGTC 2256  
  
Qy 481 GTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGAGA 540  
Db 2257 GTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGAGA 2316  
  
Qy 541 GGTTCAGATATCTTCAGGGTTGCTATGTGCGGGGCCAGTCACTCTGTGGATCTTCATG 600  
Db 2317 GGTTCAGATATCTTCAGGGTTGCTATGTGCGGGGCCAGTCACTCTGTGGATCTTCATG 2376  
  
Qy 601 ATACAGGATACCGGAACGTTATATGGGTACCCCTGACCGAGCAAGTGAACAGGGCTATTACT 660  
Db 2377 ATACAGGATACCGGAACGTTATATGGGTACCCCTGACCGAGCAAGTGAACAGGGCTATTACT 2436  
  
Qy 661 TAGGATCTGGGCCATGCAAGCAAGAAAGTTCCCTCTCGAACCAAACTGTTTACTGCTCT 720  
Db 2437 TAGGATCTGGGCCATGCAAGCAAGAAAGTTCCCTCTCGAACCAAACTGTTTACTGCTCT 2496  
  
Qy 721 TACATGTTTCTCTGGATGAGAAATGTCATTTTGGCATACATACAGTATATTTACTAGTGT 780  
Db 2497 TACATGTTTCTCTGGATGAGAAATGTCATTTTGGCATACATACAGTATATTTACTAGTGT 2556

Qy 781 TAGTGAGGCTCGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATAA 840  
Db 2557 TAGTGAGGCTCGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATAA 2616  
  
Qy 841 GAGTTCTCCTGAATCGGAGAGAACATTAATGAATGCACTTTTGGCACTACCTTCAAGAAACC 900  
Db 2617 GAGTTCTCCTGAATCGGAGAGAACATTAATGAATGCACTTTTGGCACTACCTTCAAGAAACC 2676  
  
Qy 901 TTGGATCAGCTATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG 960  
Db 2677 TTGGATCAGCTATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG 2736  
  
Qy 961 TATACACTGGCTATTTTACCAATCAGGAGGTTTAAATCAACAGAGAAACACAGAAATTCATC 1020  
Db 2737 TATACACTGGCTATTTTAAACCAATCAGGAGGTTTAAATCAACAGAGAAACACAGAAATTCATC 2796  
  
Qy 1021 ATCACATTTTGTATACCTGCAATGTAACATCTACTCTCGAAATAAATGTGTGCATGCA 1080  
Db 2797 ATCACATTTTGTATACCTGCAATGTAACATCTACTCTCGAAATAAATGTGTGCATGCA 2856  
  
Qy 1081 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAATG 1140  
Db 2857 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAATG 2916  
  
Qy 1141 ATACATATTTCTGAGAGAGCCAGCAATACCAATAGAAATTAATAAAAAA 1197  
Db 2917 ATACATATTTCTGAGAGAGCCAGCAATACCAATAGAAATTAATAAAAAA 2973  
  
RESULT 3  
ADR41222/c  
ID ADR41222 standard; cDNA; 2083 BP.  
XX ADR41222;  
XX  
XX 07-OCT-2004 (first entry)  
XX Human CD-like molecule HQDES11 cDNA, seq ID NO:21.  
XX  
XX Human; CD-like molecule; cluster of differentiation; diagnosis;  
XX prevention; immune disorder; immunodeficiency; autoimmune disorder;  
XX blood-related disorder; haematological disorder; haemostatic disorder;  
XX thrombolytic disorder; hyperproliferative disorder; cancer; tumour;  
XX apoptotic disorder; cardiovascular disorder; respiratory disorder;  
XX angiogenic disorder; neovascularisation; neurological disorder;  
XX endocrine disorder; reproductive system disorder; infectious disease;  
XX gastrointestinal disorder; drug screening; tissue regeneration;  
XX chemotaxis; gene therapy; antibody therapy; drug targeting;  
XX chromosome mapping; forensic analysis; immunophenotyping; cytostatic;  
XX haemostatic; tranquiliser; vulnery; antiinflammatory; nephrotropic;  
XX cardiant; anti-allergic; anti-HIV; antirheumatic; antiarthritic;  
XX antipruritic; immunosuppressive; vasotropic; nootropic; neuroprotective;  
XX antithyroid; thyromimetic; gynaecological; virucide; hepatotropic;  
XX antibacterial; dermatological; chromosome 15q22; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200226930-A2.  
XX  
XX 04-APR-2002.  
XX  
XX 25-SEP-2001; 2001WO-US029838.  
XX  
XX 26-SEP-2000; 2000US-0235484P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Birse CE;  
XX  
XX WPI, 2002-405050/43.  
XX  
XX P-PSDB; ADR41398.

PT Novel polynucleotides and polypeptides useful for treating, preventing or  
PT ameliorating cardiovascular, renal, neurovascular, and autoimmune  
XX disorders.  
PS Claim 4; SEQ ID NO 21; 1243pp; English.  
XX The invention relates to 167 novel human CD (cluster of differentiation)-  
CC like molecules (ADR41388-ADR41563) and to cDNAs encoding them (seqid111)-  
XX Sequence 2083 BP; 590 A; 440 C; 418 G; 635 T; 0 U; 0 Other;  
SQ

Query Match 86.9%; Score 1040; DB 7; Length 2083;  
Best Local Similarity 89.1%; Fred. No. 5.1e-293;  
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;  
QY 1 ATTTGAGGCCAACAAAGACTCCCTTTAGAGCATCCTGTAGCTAGTACGTTAGCTAA 60  
DB ATTTGAGGCCAACAAAGACTCCCTTTAGAGCATCCTGTAGCTAGTACGTTAGCTAA 1301  
QY 61 ATCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACTTCTTGTGCATCAGTC 120  
DB ATCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACTTCTTGTGCATCAGTC 1241  
QY 121 AGCACTGTGACTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180  
DB AGCACTGTGACTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 1181  
QY 181 TTTTACAAAGTATCAAGTCTGGAAGATGACCCAACTTGCACCAAGAAATTTTGGGCCA 240  
DB TTTTACAAAGTATCAAGTCTGGAAGATGACCCAACTTGCACCAAGAAATTTTGGGCCA 1121  
QY 241 CCATTTTGGATTGACAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTG 300  
DB CCATTTTGGATTGACAGGCTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTG 1061  
QY 301 AAGTACTACTGGATTACATTTATGGATGGATGCTTACAGCCCTCATGATCAGCCCTG 360  
DB AAGTACTACTGGATTACATTTATGGATGGATGCTTACAGCCCTCATGATCAGCCCTG 1001  
QY 361 GAAGAAATATCTACTGCTGCTTTCATATATGTTGCTCTCA----- 403  
DB GAAGAAATATCTACTGCTGCTTTCATATATGTTGCTCTCA----- 941  
QY 404 ----- 403  
DB 940 ATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCTAGCCTCTCTAGGTTATGNG 881  
QY 404 ----- 403  
DB 880 TTGTAGTGATAGAACACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAGCGCCTTTA 821  
QY 404 -----GGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCCAATATCTAGCTT 453  
DB 820 AATATAAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTT 761  
QY 454 CTGATATGATTTTCACTTACCTAGATCGTGTGGGCATCCACGCTGCTCTATGGAGGAT 513  
DB 760 CTGATATGATTTTCACTTACCTAGATCGTGTGGGCATCCACGCTGCTCTATGGAGGAT 701  
QY 514 ACTCTCTCTGATGGCATTAAATGACAGAGGTGAGATATCTTACGCTTGTCTATGCTGGGG 573  
DB 700 ACTCTCTCTGATGGCATTAAATGACAGAGGTGAGATATCTTACGCTTGTCTATGCTGGGG 641  
QY 574 CCCCACTCACTCTGAGGATCTTCTATGATACAGATACAGGAACTGATATATGCTGAC 633  
DB 640 CCCCACTCACTCTGAGGATCTTCTATGATACAGATACAGGAACTGATATATGCTGAC 581  
QY 634 CTGACCAAGTAAAGAGGCTTATCTAGATCTGTCGATGCTGACGAGAAAGTTC 693  
DB 580 CTGACCAAGTAAAGAGGCTTATCTAGATCTGTCGATGCTGACGAGAAAGTTC 521  
QY 694 CCTCTGAACCAATCGTTTACTGCTCTTACATGCTTCTGATGAGAAATGCCATTTTG 753

DB 520 CCTCTGAACCAATCGTTTACTCTCTTACATGGTTTCTCTGATGAGAAATGCCATTTTG 461  
QY 754 CACATACCACTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGA 813  
DB 460 CACATACCACTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGA 401  
QY 814 TCTATCTCTAGGAGACACAGCATAAAGAGTTTCTTGAATCGGGAGAACATTTGAACCTGC 873  
DB 400 TCTATCTCTAGGAGACACAGCATAAAGAGTTTCTTGAATCGGGAGAACATTTGAACCTGC 341  
QY 874 ATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGTATTGCTCTCTAAAAGTGATAT 933  
DB 340 ATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGTATTGCTCTCTAAAAGTGATAT 281  
QY 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAAAATGAGAGGTT 993  
DB 280 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAAAATGAGAGGTT 221  
QY 994 TAATCAACAGAAAACACAGAAATTTGATCATCATTTTGTATACCTGCGCATGTAACATCTAC 1053  
DB 220 TAATCAACAGAAAACACAGAAATTTGATCATCATTTTGTATACCTGCGCATGTAACATCTAC 161  
QY 1054 TCCTGAAAATAAATGTTGGTCCCATGCGGGGTCTACGGTTTGTGTAGTAAATACC 1113  
DB 160 TCCTGAAAATAAATGTTGGTCCCATGCGGGGTCTACGGTTTGTGTAGTAAATACC 101  
QY 1114 TTAACCCCATGCTCTCAAAATCAAAATGATATATCTCTGAGAGACCCAGCAATACCATA 1173  
DB 100 TTAACCCCATGCTCTCAAAATCAAAATGATATATCTCTGAGAGACCCAGCAATACCATA 41  
QY 1174 AGAATTACTAAAAA----- 1197  
DB 40 AGAATTACTAAAAA----- 17  
RESULT 4  
ID ADL13374 standard; cDNA; 2797 BP.  
XX ADL13374;  
XX 06-MAY-2004 (first entry)  
XX Human steroid-induced C3A liver cell cDNA #1103.  
XX ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;  
XX steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.  
XX Homo sapiens.  
XX US6673549-B1.  
XX 06-JAN-2004.  
XX 12-OCT-2001; 2001US-00976594.  
XX 12-OCT-2000; 2000US-0240409P.  
XX (INCY-) INCYTE CORP.  
XX Furness LM, Buchbinder JL;  
XX WPI; 2004-068610/07.  
XX Combination useful for preparing a composition for treating liver  
XX disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,  
XX comprises cDNAs that are differentially expressed in response to steroid  
XX treatment.  
XX Claim 1; SEQ ID NO 1103; 141pp; English.  
XX The invention relates to a combination comprising cDNAs that are  
XX differentially expressed in response to steroid treatment. Also included  
CC



are the following: a high throughput method for using a cDNA to detect differential expression of nucleic acids in a sample; and a high throughput method of screening molecules or compounds to identify a ligand that specifically binds a cDNA. The sample is from a subject with Wilson disease and comparison of a standard defines a stage of that disease. The high throughput method of screening molecules or compounds to identify a ligand that specifically binds a cDNA comprises: combining the combination with molecules or compounds under conditions to allow specific binding; and detecting specific binding between each cDNA and at least one molecule or compound. The molecules or compounds are regulatory proteins. The combination is useful for preparing a composition for treating liver disorders associated with steroid therapy, e.g., cirrhosis or hepatitis. The present sequence represents a human cDNA which is differentially expressed in steroid-induced C3A liver cells. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

Query Match 86.9%; Score 1040; DB 12; Length 2797;  
Best Local Similarity 89.1%; Pred. No. 5.8e-293;  
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;  
1 ATTTGAGGCCACCAAGACTCCCTTTAGACATCACCTGTAGTAGTCAGTTACGTAA 60  
1452 ATTTGAGGCCACCAAGACTCCCTTTAGACATCACCTGTAGTAGTCAGTTACGTAA 1511  
61 ATCTGAGAGGTGCAAGGCTGACGCGGTGCTCTCATCTTTCTGTCATCAGTC 120  
1512 ATCTGAGAGGTGCAAGGCTGACGCGGTGCTCTCATCTTTCTGTCATCAGTC 1571  
121 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCC 180  
1572 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCC 1631  
181 TTTCAGAGCTATCAAGTCTGAAAGATGACCACTTGGCAACAAAGAAATTTGGGCCA 240  
1632 TTTCAGAGCTATCAAGTCTGAAAGATGACCACTTGGCAACAAAGAAATTTGGGCCA 1691  
241 CCATTTTGGATTGAGAGGCTCTCTCTGCTGATATCTCTCCAGAAATTTCTCTTTTG 300  
1692 CCATTTTGGATTGAGAGGCTCTCTCTGCTGATATCTCTCCAGAAATTTCTCTTTTG 1751  
301 AAAGTACTACTGGATTATACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360  
1752 AAAGTACTACTGGATTATACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 1811  
361 GAAAGAAATATCTCTACTGTGCTTTCATATATGTTGGTGGTCTCA----- 403  
1812 GAAAGAAATATCTCTACTGTGCTTTCATATATGTTGGTGGTGGTGGTGAATA 1871  
404 ----- 403  
1872 ATCGGTTTAAAGGATCAAGATTTCCGCTTGAATACCTAGCTCTCTAGTTATGTGG 1931  
404 ----- 403  
1932 TTGTAGTGATAGACACAGGGGATCTGTGTCACGGAGGCTTAAATTTGAAGGCCCTTTA 1991  
404 -----GGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCCCAATATCTAGCTT 453  
1992 AATATAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCCCAATATCTAGCTT 2051  
454 CTCCATATGATTTTATGATCTAGATCGTGTGGGATCCAGGCTGGTCTCTATGAGGAT 513  
2052 CTCGATATGATTTTATGATCTAGATCGTGTGGGATCCAGGCTGGTCTCTATGAGGAT 2111  
514 ACCTCTCCCTGATGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTTCCTATTCCTGGGG 573  
2112 ACCTCTCCCTGATGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTTCCTATTCCTGGGG 2171  
574 CCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAAAGCTTATATGGGTCAAC 633

2172 CCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAAAGCTTATATGGGTCAAC 2231  
634 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAACAGAAAGTTTCC 693  
2232 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAACAGAAAGTTTCC 2291  
694 CCTCTGAACAAATCGTTTACTGCTCTTACATGCTTCTGATGAGATGTCATTTTG 753  
2292 CCTCTGAACAAATCGTTTACTGCTCTTACATGCTTCTGATGAGATGTCATTTTG 2351  
754 CACATACCATGATATTACTGAGTCTTCTGAGGCTGGAAGCCATATGATTTACAGA 813  
2352 CACATACCATGATATTACTGAGTCTTCTGAGGCTGGAAGCCATATGATTTACAGA 2411  
814 TCTATCCTCAGGAGACACAGCATGAAGTCTCTGAATCGGAGAACATTTGAACATGC 873  
2412 TCTATCCTCAGGAGACACAGCATGAAGTCTCTGAATCGGAGAACATTTGAACATGC 2471  
874 ATCTTTTGCATACCTTCAAGAAACCTTGGATCAGTATTGCTCTTAAAGTGATAT 933  
2472 ATCTTTTGCATACCTTCAAGAAACCTTGGATCAGTATTGCTCTTAAAGTGATAT 2531  
934 AATTTTCACCTGTGTAGACTCTCTGTATACACTGGCTATTAAACCAAATGAGGAGTT 993  
2532 AATTTTCACCTGTGTAGACTCTCTGTATACACTGGCTATTAAACCAAATGAGGAGTT 2591  
994 TAATCAACAGAAACACAGAAATGATCATCATATTGATACCTGCCATGTAACATCTAC 1053  
2592 TAATCAACAGAAACACAGAAATGATCATCATATTGATACCTGCCATGTAACATCTAC 2651  
1054 TCCTGAAAATAAATGTGGTGCATGAGGGGTCTACGGTTTGTGTAGTAACTTAATACC 1113  
2652 TCCTGAAAATAAATGTGGTGCATGAGGGGTCTACGGTTTGTGTAGTAACTTAATACC 2711  
1114 TTAACCCCACTGCTCAAAATCAAAATGATACATATTCTCTGAGAGACCAGCAATACCATA 1173  
2712 TTAACCCCACTGCTCAAAATCAAAATGATACATATTCTCTGAGAGACCAGCAATACCATA 2771  
1174 AGAATTACTAAAAAATAAATAAATAA 1197  
2772 AGAATTACTAAAAAATAAATAAATAA 2795  
RESULT 5  
AAH99934  
ID AAH99934 standard; cDNA; 3143 BP.  
XX  
AC AAH99934;  
XX  
DT 12-APR-2002 (first entry)  
XX  
DE cDNA encoding 21953 human prollyl oligopeptidase.  
XX  
KW 21953 prollyl oligopeptidase; human; proline; endopeptidase; cancer;  
cardiovascular disease; autoimmune disease; atopic allergy;  
neural disorder; vascular disorder; prostate disorder; cytostatic;  
antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;  
diabetes mellitus; arthritis; multiple sclerosis; asthma;  
Grave's disease; neuronal disorder; demyelinating disease; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 229..2877  
FT /tag= a  
FT /product= "21953 prollyl oligopeptidase"  
FT /note= "This region is specifically claimed in claim 2"  
XX  
XX WO200179473-A2.  
XX  
XX 25-OCT-2001.  
XX

PF 11-APR-2001; 2001WO-US040483.  
PR 18-APR-2000; 2000US-0197508P.  
XX (MILL-) MILLENNIUM PHARM INC.  
PA  
XX Meyers RA, Williamson M;  
PI  
XX WPI, 2002-034353/04.  
DR P-FSDB; AAG78415.  
XX  
XX New polypeptides 21953, member of human prollyl oligopeptidase family,  
PI useful as diagnostic targets and therapeutic agents for controlling  
PI cancer, lymphoma and leukemia.  
XX  
PS Claim 7; Page 100-102; 121pp; English.  
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.  
CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,  
CC antithyroid, dermatological, antipsoriatic, antiasthmatic,  
CC ophthalmological, antiinflammatory, nootropic, antiparkinsonian,  
CC anticonvulsant, gynaecological, vasotrophic, antianginal, cardiant,  
CC antiatherosclerotic, anorectic and metabolic in its action. Uses include  
CC gene therapy, expression or activity of 21953 protein modulator, it is  
CC useful for identifying a compound which binds to it and can be used in  
CC preventing, treating or detecting a cellular proliferative or  
CC differentiative disorder. The 21953 molecules can act as novel diagnostic  
CC targets and therapeutic agents for controlling disorders associated with  
CC the aberrant activity or degradation of peptide hormones e.g., disorders  
CC associated with cell differentiation and proliferation such as cancer,  
CC immune function, reproductive, neurological and cardiovascular function.  
CC The 21953 molecules are thus useful for treating and preventing cellular  
CC proliferative and differentiative disorders, hematopoietic neoplastic  
CC disorders, immune disorders such as autoimmune diseases, diabetes  
CC mellitus, arthritis, multiple sclerosis, asthma, Graves disease,  
CC neuronal disorders, demyelinating diseases, vascular disorders and  
CC metabolism or pain disorders, this sequence represents the cDNA encoding  
CC sequence of 21953 human prollyl oligopeptidase  
XX  
SQ Sequence 3143 BP; 943 A; 644 C; 712 G; 844 T; 0 U; 0 Other;

Query Match 86.9%; Score 1040; DB 6; Length 3143;  
Best Local Similarity 89.1%; Pred. No. 6.1e-293;  
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

Qy 1 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAA 60  
Db 1796 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAA 1855

Qy 61 ATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTC 120  
Db 1856 ATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTC 1915

Qy 121 AGCACTGTGACTTCTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180  
Db 1916 AGCACTGTGACTTCTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 1975

Qy 181 TTTACAGCTATCAAGCTCTGAAGATGACCCAACTTGCAAAACAAGGAAATTTGGGCCA 240  
Db 1976 TTTACAGCTATCAAGCTCTGAAGATGACCCAACTTGCAAAACAAGGAAATTTGGGCCA 2035

Qy 241 CCATTTGGATTACAGAGGCTCTCTCTGCTGCTATCTCTGCTGCTGCTGCTGCTGCTGCT 300  
Db 2036 CCATTTGGATTACAGAGGCTCTCTCTGCTGCTATCTCTGCTGCTGCTGCTGCTGCTGCT 2095

Qy 301 AAGTACTACTGATTTACATTTGATGGATGCTTACAGCCCTCATGATCTACAGCTG 360  
Db 2096 AAGTACTACTGATTTACATTTGATGGATGCTTACAGCCCTCATGATCTACAGCCCTG 2155

Qy 361 GAAAGAAATATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403  
Db 2156 GAAAGAAATATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2215

Qy 404 ----- 403  
Db 2216 ATCGTTTAAAGGAGTCAGTATTTCGGCTTGAATACCTAGCCCTCTCTAGGTATTATGTGG 2275  
Qy 404 ----- 403  
Db 2276 TTGTAGTAGTAGACAACAGGGGATTCCTGTCA CCGAGGGCTTAAATTTGAAAGGCGCTTTA 2335  
Qy 404 -----GGGTCAAATAGAAAATTTGACGATCAGGTGGAAGGACTCAATATCTAGCTT 453  
Db 2336 AATAATAAAATGGGTCAAATAGAAAATTTGACGATCAGGTGGAAGGACTCAATATCTAGCTT 2395  
Qy 454 CTCGATATGATTTTCATTGACTTAGATCGTGTGGGATCCACGGCTGGTCCCTATGAGAGAT 513  
Db 2396 CTCGATATGATTTTCATTGACTTAGATCGTGTGGGATCCACGGCTGGTCCCTATGAGAGAT 2455  
Qy 514 ACCTCTCCCTGATGGGCAATTAATGACAGGTCAGATATCTTCAGGGTTCCTATTGCTGGGG 573  
Db 2456 ACCTCTCCCTGATGGGCAATTAATGACAGGTCAGATATCTTCAGGGTTCCTATTGCTGGGG 2515  
Qy 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATGGGTCAAC 633  
Db 2516 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATGGGTCAAC 2575  
Qy 634 CTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGCCATCGCAACGACAAAGTTC 693  
Db 2576 CTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGCCATCGCAACGACAAAGTTC 2635  
Qy 694 CCTCTGAACCAATCGTTTACTGCTTACATGCTTCTGATGAGGATGCTCCATTTTG 753  
Db 2636 CCTCTGAACCAATCGTTTACTGCTTACATGCTTCTGATGAGGATGCTCCATTTTG 2695  
Qy 754 CACATACCATATATTACTGAGTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGA 813  
Db 2696 CACATACCATATATTACTGAGTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGA 2755  
Qy 814 TCTATCTCTAGGAGACACAGATAAGATTCCTGAAATCGGGAGAACATTTATGAACTGC 873  
Db 2756 TCTATCTCTAGGAGACACAGATAAGATTCCTGAAATCGGGAGAACATTTATGAACTGC 2815  
Qy 874 ATCTTTTGCACTACCTTCAAGAAAACCTTTGGATCACTGATTTGCTCTCTAAAAGTGATAT 933  
Db 2816 ATCTTTTGCACTACCTTCAAGAAAACCTTTGGATCACTGATTTGCTCTCTAAAAGTGATAT 2875  
Qy 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTT 993  
Db 2876 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTT 2935  
Qy 994 TAATCAACAGAAAACACAGAAATTTGATCATCAATTTTGATACCTGCCATGTAACTCTAC 1053  
Db 2936 TAATCAACAGAAAACACAGAAATTTGATCATCAATTTTGATACCTGCCATGTAACTCTAC 2995  
Qy 1054 TCCTGAAAATAAATGTGGTGCATCGAGGGGTCTACGGTTTGTGTAGTAACTTAATACC 1113  
Db 2996 TCCTGAAAATAAATGTGGTGCATCGAGGGGTCTACGGTTTGTGTAGTAACTTAATACC 3055  
Qy 1114 TTAACCCCATGCTCCAAATCAATGATACATATTTCTGAGAGACCCAGCAATACCATA 1173  
Db 3056 TTAACCCCATGCTCCAAATCAATGATACATATTTCTGAGAGACCCAGCAATACCATA 3115  
Qy 1174 AGAATTTACTAAAAA 1197  
Db 3116 AGAATTTACTAAAAA 3139

RESULT 6  
AAC85694  
ID AAC85694 standard; cDNA; 3120 BP.  
XX  
AC AAC85694;  
XX  
DT 29-JUN-2001 (first entry)  
XX



```
QY 1174 AGAATTACTAAAAA 1193
DB 3101 AGAATTACTAAAAA 3120

RESULT 7
ID AAD38956
AC AAD38956;
DT 23-SEP-2002 (first entry)
DE Human dipeptidyl peptidase 8 (DPP8) cDNA.
KW Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW antiviral; enzyme; gene; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 214..2862
FT /tag= a
FT /product= "Human DPP8 protein"
FN WO200234900-A1.
PD 02-MAY-2002.
PP 29-OCT-2001; 2001WO-AU001388.
PR 27-OCT-2000; 2000AU-00001078.
PA (UNSY ) UNIV SYDNEY.
PI Abbott CA, Gorrell MD;
XX WPI; 2002-454646/48.
DR P-PSDB; AAE24170.
XX
XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
XX of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
XX type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
XX infection.
XX Example; Fig 1; 91pp; English.
XX
XX The present invention relates to dipeptidyl peptidase (DPP) proteins and
XX polynucleotides encoding such proteins. The DPP peptides are useful for
XX screening inhibitors of DPP catalytic activity. The inhibitors are useful
XX for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
XX rejection and HIV (human immuno deficiency virus) infection. The present
XX sequence is human DPP8 cDNA
XX
SQ Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;

Query Match 86.5%; Score 1036; DB 6; Length 3120;
Best Local Similarity 89.0%; Pred. No. 9e-292;
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAA 60
DB 1781 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAA 1840
QY 61 ATCTCGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTCTTGTGTCATCAGTC 120
DB 1841 ATCTCGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTCTTGTGTCATCAGTC 1900
QY 121 AGCACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCC 180
DB 1901 AGCACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCC 1960
```

```
QY 181 TTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGGCCA 240
DB 1961 TTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGGCCA 2020
QY 241 CCATTTTGGATTCAGACAGGTCCTCTTCTGATATATCTCTCCAGAAAATTTTCTCTTTTG 300
DB 2021 CCATTTTGGATTCAGACAGGTCCTCTTCTGATATATCTCTCCAGAAAATTTTCTCTTTTG 2080
QY 301 AAAGTACTACTGGATTACATTTGATGGGATCTCTCAAGGCTCATGATCTACAGCCTG 360
DB 2081 AAAGTACTACTGGATTACATTTGATGGGATCTCTCAAGGCTCATGATCTACAGCCTG 2140
QY 361 GAAAGAAAATATCCTACTGTCTGTTTATATATGGTGGTCTCA----- 403
DB 2141 GAAAGAAAATATCCTACTGTCTGTTTATATATGGTGGTCTCTCAGGTGCAAGTTGGTGAATA 2200
QY 404 ----- 403
DB 2201 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGSCCTCTCTAGGTTATGTGG 2260
QY 404 ----- 403
DB 2261 TTGTAGTGATAGACAAACAGGGGATCCTGTACCAGGGGCTTAAATTTTGAAGGCGCTTTA 2320
QY 404 -----GGGTCAAATAGAAAATTTGACGATCAGGTGSAAGGACTCCAATATCTAGCTT 453
DB 2321 AATATAAAATGGGTCAAATAGAAAATTTGACGATCAGGTGSAAGGACTCCAATATCTAGCTT 2380
QY 454 CTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCAGGCTGCTCTATCGAGAT 513
DB 2381 CTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCAGGCTGCTCTATCGAGAT 2440
QY 514 ACCTCTCCGTGATGGCAATTAATGACAGGTGAGATATCTTCAGGTTGCTATTCCTGGGG 573
DB 2441 ACCTCTCCGTGATGGCAATTAATGACAGGTGAGATATCTTCAGGTTGCTATTCCTGGGG 2500
QY 574 CCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC 633
DB 2501 CCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC 2560
QY 634 CTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATCGAAGCAAGAAAGTTCC 693
DB 2561 CTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATCGAAGCAAGAAAGTTCC 2620
QY 694 CCTCTGAACCAAATCGTTTACTGTCTTTACATGTTTCTCTGGATGAGAATGTCCATTTTG 753
DB 2621 CCTCTGAACCAAATCGTTTACTGTCTTTACATGTTTCTCTGGATGAGAATGTCCATTTTG 2680
QY 754 CACATACCAAGTATATTACTGAGTTTCTTAGTGAGGCTGGAAAGCCATATGATTTACAGA 813
DB 2681 CACATACCAAGTATATTACTGAGTTTCTTAGTGAGGCTGGAAAGCCATATGATTTACAGA 2740
QY 814 TCTATCCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAAACATTATGAACCTGC 873
DB 2741 TCTATCCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAAACATTATGAACCTGC 2800
QY 874 ATCTTTTGCACTACCTTCAAGAAAACCTTTGATCAGCTATTGTGCTCTAAAAGTATAT 933
DB 2801 ATCTTTTGCACTACCTTCAAGAAAACCTTTGATCAGCTATTGTGCTCTAAAAGTATAT 2860
QY 934 AATTTTGACCTCTGTAGAACTCTCTGGTATACACCTGCTATTTAAACCAATGAGGAGTT 993
DB 2861 AATTTTGACCTCTGTAGAACTCTCTGGTATACACCTGCTATTTAAACCAATGAGGAGTT 2920
QY 994 TAATCAACAGAAACACAGAAATGATCATACATTTTGTATGATCCTGCAATACATCTAC 1053
DB 2921 TAATCAACAGAAACACAGAAATGATCATACATTTTGTATGATCCTGCAATACATCTAC 2980
QY 1054 TCTGTAAATTAATCTGTGTGATGCGGGGTCTACCGTTTGTGTGTGTGTGTGTGTGTGTGT 1113
DB 2981 TCTGTAAATTAATCTGTGTGATGCGGGGTCTACCGTTTGTGTGTGTGTGTGTGTGTGTGT 3040
```



Qy	992	TTTAAATCAACAGAAAAACA	CAGAATTTGATCATCACTTTGATACCTGCGCATGTAACTCT	1051
Db	2401	TTTAAATCAACAGAAAAACA	CAGAATTTGATCATCACTTTGATACCTGCGCATGTAACTCT	2460
Qy	1052	ACTCTGAAAAATTAATGTGGTGCCATCAGGGGTCTACGGTTTGTGGTAGTAACTCTTAATA	1111	
Db	2461	ACTCTGAAAAATTAATGTGGTGCCATCAGGGGTCTACGGTTTGTGGTAGTAACTCTTAATA	2520	
Qy	1112	CCTTAAACCCCATGCTCAAAATCAATGATACATATCTCTGAGAGACCCAGCAATACCA	1171	
Db	2521	CCTTAAACCCCATGCTCAAAATCAATGATACATATCTCTGAGAGACCCAGCAATACCA	2580	
Qy	1172	TAAGAAATTACTAAAAA	1197	
Db	2581	TAAGAAATTACTAAAAA	2606	

## RESULT 9

ABK83327	ABK83327 standard; CDNA; 4829 BP.
ID	XX
XX	ABK83327
AC	XX
XX	XX
DT	12-AUG-2002 (first entry)
XX	XX
DE	CDNA encoding human DPRP-1 splice variant #3.
XX	XX
KW	Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW	DPRPV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW	heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW	ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW	dyskinesia; reproductive disorder; inflammatory disorder;
KW	metabolic disorder; gene; es.

XX	SQ	Sequence 4829 BP; 1466 A; 896 C; 1017 G; 1460 T; 0 U; 0 Other;
		Query Match 86.3%; Score 1033.2; DB 6; Length 4829;
		Best Local Similarity 88.7%; Pred. No. 7.3e-291;
		Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;
QY	1	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 60
DB	1781	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 1840
QY	61	ATCCTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATCTCTGTCGATCAGTC 120
DB	1841	ATCCTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATCTCTGTCGATCAGTC 1900
QY	121	AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACAGAGGAATCCACACTGTGTGCC 180
DB	1901	AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACAGAGGAATCCACACTGTGTGCC 1960
QY	181	TTTACAGCTCATCAAGTCCGTGAGATGACCAACTTGGCAAAACAAGGAATTTTGGCCA 240
DB	1961	TTTACAGCTCATCAAGTCCGTGAGATGACCAACTTGGCAAAACAAGGAATTTTGGCCA 2020
QY	241	CCATTTGGATTCAGCAGGCTCTTCTCGATATATCTCTCCAGAAATTTTCTCTTTTG 300
DB	2021	CCATTTGGATTCAGCAGGCTCTTCTCGATATATCTCTCCAGAAATTTTCTCTTTTG 2080
QY	301	AAAGTACTACTGGATTATCATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
DB	2081	AAAGTACTACTGGATTATCATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 2140
QY	361	GAAGAAATATCTACTGTCTGTTTATATATGGTGGTCCCTC----- 402
DB	2141	GAAGAAATATCTACTGTCTGTTTATATATGGTGGTCTCCTCAGGTGCAGTTGGTGAA 2200
QY	403	----- 402
DB	2201	TAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGT 2260
QY	403	----- 402
DB	2261	GTTGTGTAGTATAGACAAACAGGGGATCCTGTCAACCGGGCTTAAATTTGAAGCGCCTT 2320
QY	403	-----AGGGTCAAATAGAAATTTGACGATCAGGTGGAAGACTCCAAATATCTAGC 451
DB	2321	TAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGACTCCAAATATCTAGC 2380
QY	452	TTCTCGATATGATTTTCATTGACTTAGATCGGTGGGCATCCACGGCTGGTCCCTATGGAG 511
DB	2381	TTCTCGATATGATTTTCATTGACTTAGATCGGTGGGCATCCACGGCTGGTCCCTATGGAG 2440
QY	512	ATACCTCTCCGTAGGCAATTAATGCAAGCTCAGATATCTTCAAGGTTGCTATGCTGG 571
DB	2441	ATACCTCTCCGTAGGCAATTAATGCAAGCTCAGATATCTTCAAGGTTGCTATGCTGG 2500
QY	572	GGCCCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAAAGTTATATGGGTCA 631
DB	2501	GGCCCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAAAGTTATATGGGTCA 2560
QY	632	CCCTGACCAAGATGAACAGGGCTATTTCTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTT 691
DB	2561	CCCTGACCAAGATGAACAGGGCTATTTCTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTT 2620
QY	692	CCCCCTCTGAAACCAATCGTTTACTGCTCTTATCATGTTTCTCTGGATGAGAATGTCATTT 751
DB	2621	CCCCCTCTGAAACCAATCGTTTACTGCTCTTATCATGTTTCTCTGGATGAGAATGTCATTT 2680
QY	752	TGCACATACCAAGTATATTACTAGTTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACA 811
DB	2681	TGCACATACCAAGTATATTACTAGTTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACA 2740
QY	812	GATCTATCCTCAGGAGAGACACAGCATAAAGAGTTTCTTGAATCGGAGAAATTATGAACT 871



Db 2741 GATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTATGAAC 2800  
Qy 872 GCATCTTTTGACACTACCTTCAGAAAAACCTTCGATCAGGTATTCGTCTCTAAGAGTAT 931  
Db 2801 GCATCTTTTGACACTACCTTCAGAAAAACCTTCGATCAGGTATTCGTCTCTAAGAGTAT 2860  
Qy 932 ATAAATTTGACCTGTGTAGAACTCTCTGGTATATACACTGCTATTTAACCAATGAGGAG 991  
Db 2861 ATAAATTTGACCTGTGTAGAACTCTCTGGTATATACACTGCTATTTAACCAATGAGGAG 2920  
Qy 992 TTTAATCAACAGAAAAACACAGAAATGATCATCATCTTTTGATACCTGCGCATGTAACTCT 1051  
Db 2921 TTTAATCAACAGAAAAACACAGAAATGATCATCATCTTTTGATACCTGCGCATGTAACTCT 2980  
Qy 1052 ACTCTGAAATAAATGTGTGTCATGCGAGGGTCTACGGTTTGTGTAGTAATCTTAATA 1111  
Db 2981 ACTCTGAAATAAATGTGTGTCATGCGAGGGTCTACGGTTTGTGTAGTAATCTTAATA 3040  
Qy 1112 CTTAACCACCATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 1171  
Db 3041 CTTAACCACCATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCA 3100  
Qy 1172 TAAGATTACTTAAAAAATAAAAAA 1197  
Db 3101 TAAGATTACTTAAAAAATAAAAAA 3126

RESULT 10

ABK12892  
ID ABK12892 standard; cDNA; 3106 BP.  
AC ABK12892;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Human protéase PRTS-9 cDNA sequence.  
XX  
KW Human; protéase; PRTS; gastrointestinal; Crohn's disease; cancer;  
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
KW cell proliferative disorder; developmental disorder; epilepsy;  
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
KW reproductive disorder; endometriosis; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 203...2851  
FT /\*tag= a  
FT /product= "Human protéase PRTS-9"

XX  
PN WO200198468-A2.  
XX  
PD 27-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019178.  
XX  
PR 16-JUN-2000; 2000US-0212336P.  
PR 22-JUN-2000; 2000US-0213955P.  
PR 29-JUN-2000; 2000US-0215396P.  
PR 07-JUL-2000; 2000US-0215821P.  
PR 14-JUL-2000; 2000US-0218946P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;  
PI Delegeane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;  
PI Azimzal Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
PI Kallick DA;  
XX  
DR WPI; 2002-090437/12.  
DR P-PSDB; AAU74749.

XX  
PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in  
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative  
PT (e.g. cancer) disorders.  
XX  
PS Claim 5; Page 166-167; 177pp; English.  
XX  
CC The present invention relates to twenty one new human proteases, referred  
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the  
CC invention are useful in the diagnosis, treatment and prevention of  
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's  
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial  
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency  
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.  
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,  
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's  
CC disease and reproductive e.g. infertility and endometriosis disorders.  
CC Numerous other examples of each disorder are given in the specification.  
CC The present nucleic acid sequence encodes the human protease PRTS-9  
CC protein of the invention  
XX  
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;  
Query Match 86.3%; Score 1033; DB 6; Length 3106;  
Best Local Similarity 89.0%; Pred. No. 6.8e-291;  
Matches 1190; Conservative 0; Mismatches 0; Indels 147; Gaps 1;  
Qy 1 ATTTTGAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAA 60  
Db 1770 ATTTTGAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 1829  
Qy 61 ATCTCGAGAGGTGACAAAGGCTGACCTGACCGTGGTACTCACATTTCTTGTGCATCAGTC 120  
Db 1830 ATCTCGAGAGGTGACAAAGGCTGACCTGACCGTGGTACTCACATTTCTTGTGCATCAGTC 1889  
Qy 121 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACATGTGTGTC 180  
Db 1890 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACATGTGTGTC 1949  
Qy 181 TTTTCAAGCTATCAAGTCTGTAAGATGACCCAACTTTGCAAAACAAGAAATTTTGGGCA 240  
Db 1950 TTTTCAAGCTATCAAGTCTGTAAGATGACCCAACTTTGCAAAACAAGAAATTTTGGGCA 2009  
Qy 241 CCATTTTGGATTGACGAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTG 300  
Db 2010 CCATTTTGGATTGACGAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTG 2069  
Qy 301 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTG 360  
Db 2070 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAGGCTCATGATCTACAGCTG 2129  
Qy 361 GAAAGAAATATCTTACTGTCTGTTTATATGTTGGTGGTCTCA----- 403  
Db 2130 GAAAGAAATATCTTACTGTCTGTTTATATGTTGGTGGTCTCAAGGTGAGTTGGTAATA 2189  
Qy 404 ----- 403  
Db 2190 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTATTATGTG 2249  
Qy 404 ----- 403  
Db 2250 TTGTAGTGATAGACACAGGGGATCTCTGTCACCGAGGGCTTAAATTTGAAGCGCTTTA 2309  
Qy 404 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGCTCCCAATATCTAGCTT 453  
Db 2310 AATAATAAATGGGTCAAATAGAAATTCAGATCAGGTGGAAGGCTCCCAATATCTAGCTT 2369  
Qy 454 CTCGATGATGATTTTCATTGACTTGTAGTGGGATCCACCGCTGGTCTCTATGAGGAT 513  
Db 2370 CTCGATGATGATTTCAITGACTTGTAGTGGGATCCACCGCTGGTCTCTATGAGGAT 2429  
Qy 514 ACCTCTCCCTGATGGCAATTAATGAGAGGTACAGATATCTTCAGGGTGTGCTATTGTG 573



Db 2430 ACCCTCCCTGCTGATGATTAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGG 2489  
Qy 574 CCCAGTCACCTCTGTCGATCTCTATGATACAGGATACACGGAAGTTATATGGTCAAC 633  
Db 2490 CCCAGTCACCTCTGTCGATCTCTATGATACAGGATACACGGAAGTTATATGGTCAAC 2549  
Qy 634 CTGACGAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCC 693  
Db 2550 CTGACGAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCC 2609  
Qy 694 CCTCTGAACCAATCGTTTACTGCTCTTACATGGTTCTCTGATGAGAAATGTCATTTTG 753  
Db 2610 CCTCTGAACCAATCGTTTACTGCTCTTACATGGTTCTCTGATGAGAAATGTCATTTTG 2669  
Qy 754 CACATACCACTATATCTAGGTTTCTTATAGTGGGCTGGAAAGCCATATGATTTACAGA 813  
Db 2670 CACATACCACTATATCTAGGTTTCTTATAGTGGGCTGGAAAGCCATATGATTTACAGA 2729  
Qy 814 TCTATCTCTCAGGAGACACAGCATAAAGATTCCTGAAATCGGAGAACATTTATGAATGC 873  
Db 2730 TCTATCTCTCAGGAGACACAGCATAAAGATTCCTGAAATCGGAGAACATTTATGAATGC 2789  
Qy 874 ATCTTTTGGCACTACCTTCAAGAAACCTTGGATCACGTATTGCTGCTTAAAGTGTAT 933  
Db 2790 ATCTTTTGGCACTACCTTCAAGAAACCTTGGATCACGTATTGCTGCTTAAAGTGTAT 2849  
Qy 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAATGAGGAGTT 993  
Db 2850 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTAACCAATGAGGAGTT 2909  
Qy 994 TATCATCAGAAACACAGAAATGATCATCATATTTGATCATCGGCTGATGATCATCTAC 1053  
Db 2910 TATCATCAGAAACACAGAAATGATCATCATATTTGATCATCGGCTGATGATCATCTAC 2969  
Qy 1054 TCTGAAATAAATGTGTGTCATGACAGGGCTACAGGTTTGTGGTATCATCTATAC 1113  
Db 2970 TCTGAAATAAATGTGTGTCATGACAGGGCTACAGGTTTGTGGTATCATCTATAC 3029  
Qy 1114 TTAACCCCATCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCAT 1173  
Db 3030 TTAACCCCATCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCAT 3089  
Qy 1174 AGAATTACTAAAAAAA 1190  
Db 3090 AGAATTACTAAAAAAA 3106

RESULT 11  
ACA92425  
ID ACA92425 standard; DNA; 2929 BP.  
XX AC ACA92425;  
XX DT 15-JUL-2003 (first entry)  
XX DB DNA encoding human PMMM-10.  
XX KW Human; protein modification and maintenance molecule; PMMM; cancer;  
KW cell proliferation disorder; atherosclerosis; neurological disorder;  
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;  
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;  
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;  
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;  
KW antiinflammatory; thyromimetic; gene; ds.  
OS Homo sapiens.  
XX PN W02003031939-A2.  
XX PD 17-APR-2003.  
XX

PP 11-OCT-2002; 2002WO-US032850.  
XX 12-OCT-2001; 2001US-0329689P.  
PR 25-OCT-2001; 2001US-0335703P.  
PR 09-NOV-2001; 2001US-0348887P.  
PR 28-NOV-2001; 2001US-0334145P.  
PR 06-DEC-2001; 2001US-0337451P.  
PR 14-DEC-2001; 2001US-0340584P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
PI Tran UK, Becha SD, Duggan BM, Lee BA, Griffin JA, Li JX;  
PI Sprague WW, Hafalla AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;  
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;  
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
DR WPI; 2003-430274/40.  
DR P-PSDB; ABU92030.  
XX New human protein modification and maintenance molecules (PMMM), useful  
PT for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or  
PT infections.  
XX Claim 5; Page 289; 311pp; English.  
XX The present invention relates to the isolation of human protein  
CC modification and maintenance molecules (PMMM), and the polynucleotide  
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM  
CC 1 to PMMM-40) are disclosed. The sequences of the invention are useful  
CC for diagnosing a condition or disease associated with the expression of  
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and  
CC generating an expression profile of a sample containing the  
CC polynucleotides. The diseases or conditions associated with decreased  
CC expression or overexpression of PMMM are cell proliferation disorders  
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,  
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
CC syndrome), gastrointestinal or epithelial disorders, and infections. The  
CC PMMM polypeptides or their fragments are useful in screening compounds  
CC for effectiveness as agonists or antagonists of the polypeptides, or in  
CC altering the expression of the target polynucleotide and compounds that  
CC specifically bind to, or modulate the activity of the polypeptide.  
CC ACA92416-ACA92455 encode the human PMMM polypeptides of the invention  
XX  
SQ Sequence 2929 BP; 864 A; 606 C; 661 G; 798 T; 0 U; 0 Other;  
Query Match 86.0%; Score 1030; DB 10; Length 2929;  
Best Local Similarity 89.0%; Pred. No. 5e-290;  
Matches 1187; Conservative 0; Mismatches 0; Indels 147; Gaps 1;  
Qy 1 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTGTAGTGTAGTGTAA 60  
Db 1596 ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTGTAGTGTAGTGTAA 1655  
Qy 61 ATCTCGAGAGGTGACNAGGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
Db 1656 ATCTCGAGAGGTGACNAGGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1715  
Qy 121 AGACCTGTGACTCTTTTAAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTAT 180  
Db 1716 AGACCTGTGACTCTTTTAAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTAT 1775  
Qy 181 TTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAAAGAAATTTTGGGCCA 240  
Db 1776 TTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAAAGAAATTTTGGGCCA 1835  
Qy 241 CCATTTTGGATTACAGAGTCTCTTCTCTGATATATCTCTCCAGAAATTTTCTCTTTTG 300  
Db 1836 CCATTTTGGATTACAGAGTCTCTTCTCTGATATATCTCTCCAGAAATTTTCTCTTTTG 1895  
Qy 301 AAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTG 360

Db 1896 AAAGTACTCTGGATTTACATTTGATGGATGCTCTACAAAGCCTCATGATCTACAGCGTG 1955  
Qy 361 GAAAGAAATATCCTACTGCTGCTTATATATATGTTGCTCTCA----- 403  
Db 1956 GAAAGAAATATCCTACTGCTGCTTATATATATGTTGCTCTCAAGTGCGAGTGGTGAATA 2015  
Qy 404 ----- 403  
Db 2016 ATCGGTTTAAAGGNGTCAAGTATTTCCGGCTTGATACCTAGCTCTCTAGTTATGTGG 2075  
Qy 404 ----- 403  
Db 2076 TTGTAGTATAGACAACAGGGGATCTGTACCGAGGGCTTAAATTTGAAGCGCGCTTTA 2135  
Qy 404 -----GGGTCAAAATGAATTTGAGCATCAGGTGGAAGGACTCCAAATATCTAGCTT 453  
Db 2136 AATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTT 2195  
Qy 454 CTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCAGCGCTGCTCTATGGAGAT 513  
Db 2196 CTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCAGCGCTGCTCTATGGAGAT 2255  
Qy 514 ACTCTCCCTGATGGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTTCCTATTGCTGGG 573  
Db 2256 ACTCTCCCTGATGGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTTCCTATTGCTGGG 2315  
Qy 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGTTCATATGCTGTCACC 633  
Db 2316 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGTTCATATGCTGTCACC 2375  
Qy 634 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 693  
Db 2376 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 2435  
Qy 694 CCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTG 753  
Db 2436 CCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTG 2495  
Qy 754 CACATACAGTATATCTAGTATTTTATGAGGGCTGGAAGCCATATGATTTACAGA 813  
Db 2496 CACATACAGTATATCTAGTATTTTATGAGGGCTGGAAGCCATATGATTTACAGA 2555  
Qy 814 TGTATCTCAGGAGACACAGCATAGAGTTTCTGAATCGGAGAACATATCAACTGC 873  
Db 2556 TGTATCTCAGGAGACACAGCATAGAGTTTCTGAATCGGAGAACATATCAACTGC 2615  
Qy 874 ATCTTTTGCACTACCTTCAAGAAACCTTTGGATCAGTATTTGCTCTTAAAGTGATAT 933  
Db 2616 ATCTTTTGCACTACCTTCAAGAAACCTTTGGATCAGTATTTGCTCTTAAAGTGATAT 2675  
Qy 934 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTATTAACCAAAATGAGAGGTT 993  
Db 2676 AATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTATTAACCAAAATGAGAGGTT 2735  
Qy 994 TAATCAACAGAAACACAGAAATGATCATCAATTTGATACCTGCCATGTAAACATCTAC 1053  
Db 2736 TAATCAACAGAAACACAGAAATGATCATCAATTTGATACCTGCCATGTAAACATCTAC 2795  
Qy 1054 TCTGAAAAATAATGTGGTCCATGACGGGGCTACGGGTTGTGGTAGTAATCTAATPACC 1113  
Db 2796 TCTGAAAAATAATGTGGTCCATGACGGGGCTACGGGTTGTGGTAGTAATCTAATPACC 2855  
Qy 1114 TTAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCAGCAATACCATTA 1173  
Db 2856 TTAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCAGCAATACCATTA 2915  
Qy 1174 AGAATTACTAAAA 1187  
Db 2916 AGAATTACTAAAA 2929

RESULT 12

ABV76411  
ID ABV76411 standard; cDNA; 1821 BP.  
XX AC ABV76411;  
XX AC ABV76411;  
DT 07-FEB-2003 (first entry)  
XX DE Dipeptidyl peptidase IV protein 27.5-encoding cDNA.  
XX DE Dipeptidyl peptidase IV protein 27.5; recombinant production;  
KW Stenotrophomonas maltophilia dipeptidyl peptidase IV homologue;  
KW gene therapy; diabetes; tumour; cancer; cytostatic; gene; ss.  
XX OS Unidentified.  
XX FH Key Location/Qualifiers  
FT CDS 826..1578  
FT /\*tag= a  
FT /product= "Dipeptidyl peptidase IV protein 27.5"  
XX CN1342768-A.  
XX PD 03-APR-2002.  
XX PF 12-SEP-2000; 2000CN-00125127.  
XX PR 12-SEP-2000; 2000CN-00125127.  
XX PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.  
XX PI Mao Y, Xie Y;  
XX DR WPI: 2002-529809/57.  
XX DR P-PSDB; ABB99949.  
XX PT New dipeptidyl peptidase IV protein 27.5 and encoding polynucleotide,  
XX useful for treating diabetes and tumours.  
XX PS Claim 6; Page 24-25 (Disclosure); 32pp; Chinese.  
XX CC The invention relates to dipeptidyl peptidase IV protein 27.5 (ABB99949)  
CC and nucleic acids encoding it (ABV76411). The protein has a molecular  
CC weight of 27.5 kD and has 40% identity and 58% homology over a 227 amino  
CC acid stretch with Stenotrophomonas maltophilia dipeptidyl peptidase IV  
CC (Genbank accession number D83263). The invention also relates to a method  
CC for the recombinant production of the protein, an antagonist of the  
CC protein, and the use of the protein, gene and antagonist in therapeutic  
CC applications. Dipeptidyl peptidase IV protein 27.5 can be used in the  
CC treatment of a variety of diseases such as diabetes and tumours. The  
CC present sequence represents cDNA encoding dipeptidyl peptidase IV protein  
CC 27.5  
XX SQ Sequence 1821 BP; 533 A; 374 C; 391 G; 523 T; 0 U; 0 Other;  
Query Match 85.3%; Score 1021; DB 6; Length 1821;  
Best Local Similarity 88.9%; Pred. No; 1.7e-287;  
Matches 1178; Conservative 0; Mismatches 0; Indels 147; Gaps 1;  
Qy 1 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGTAA 60  
Db 497 ATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGTAA 556  
Qy 61 ATCTGAGAGGTGACAAAGGCTGACGCGGTGGTACTCACATTTCTGTCATCAGTC 120  
Db 557 ATCTGAGAGGTGACAAAGGCTGACGCGGTGGTACTCACATTTCTGTCATCAGTC 616  
Qy 121 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180  
Db 617 AGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 676  
Qy 181 TTTTCAAGCTATCAAGTCTTGAAGATGACCCAACTTGCACAAAGAAATTTTGGGCA 240  
Db 677 TTTTCAAGCTATCAAGTCTTGAAGATGACCCAACTTGCACAAAGAAATTTTGGGCA 736

```
QY 241 CCATTTTGGATTGAGGAGTCTCTTCTGACTATACCTCTCCAGAAATTTTCTCTTTTG 300
DB 737 CCATTTTGGATTGAGGAGTCTCTTCTGACTATACCTCTCCAGAAATTTTCTCTTTTG 796
QY 301 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
DB 797 AAAGTACTACTGGATTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG 856
QY 361 GAAAGAAATATCCTACTGTGCTGTTTATATATATGTTGGTCTCTCA----- 403
DB 857 GAAAGAAATATCCTACTGTGCTGTTTATATATATGTTGGTCTCTCAGGTGAGTTGGTAATA 916
QY 404 ----- 403
DB 917 ATCGGTTTTAAAGGATCAAGATTATTCGGCTTGAATACCTAGCTCTCTAGTTTATGTGG 976
QY 404 ----- 403
DB 977 TTGTAGTATAGACACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGGCGCCTTTA 1036
QY 404 -----GGCTCAATAGAAATGACGATCAGGTGAGGAGTCCATATATCTAGCTT 453
DB 1037 AATATAAATGGGTCAATAGAAATGACGATCAGGTGAGGAGTCCATATATCTAGCTT 1096
QY 454 CTCGATATGATTTTCAATGATCTAGATCGTGTGGGATCCAGCGCTGGTCTTATGGAGGAT 513
DB 1097 CTCGATATGATTTTCAATGATCTAGATCGTGTGGGATCCAGCGCTGGTCTTATGGAGGAT 1156
QY 514 ACCTCTCCCTGATGGCAATTAATGAGAGTCAAGATATCTTCAAGGGTTGCTATGCTGGGG 573
DB 1157 ACCTCTCCCTGATGGCAATTAATGAGAGTCAAGATATCTTCAAGGGTTGCTATGCTGGGG 1216
QY 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC 633
DB 1217 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAAC 1276
QY 634 CTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 693
DB 1277 CTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCC 1336
QY 694 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTG 753
DB 1337 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTG 1396
QY 754 CACATACAGTATATTACTGAGTGTGTTTGTAGTGGGCTGGAAGCCATATGATTTACAGA 813
DB 1397 CACATACAGTATATTACTGAGTGTGTTTGTAGTGGGCTGGAAGCCATATGATTTACAGA 1456
QY 814 TCTATCCTCAGGAGACACGATAGAGTTCTGTAATCGGAGACATATGCACTCC 873
DB 1457 TCTATCCTCAGGAGACACGATAGAGTTCTGTAATCGGAGACATATGCACTCC 1516
QY 874 ATCTTTTGAATACCTTCAAGAAACCTTGGATCAAGTATGCTGCTCTAAAGTGTATAT 933
DB 1517 ATCTTTTGAATACCTTCAAGAAACCTTGGATCAAGTATGCTGCTCTAAAGTGTATAT 1576
QY 934 AATTTTGACCTGTGTAGAACTCTCTGTTATACACTGGCTATTTAACCAATGAGGAGTT 993
DB 1577 AATTTTGACCTGTGTAGAACTCTCTGTTATACACTGGCTATTTAACCAATGAGGAGTT 1636
QY 994 TAATCAACAGAAACACAGAAATGATCATCATTTTGAATCTGCTGCCATGTAACATCTAC 1053
DB 1637 TAATCAACAGAAACACAGAAATGATCATCATTTTGAATCTGCTGCCATGTAACATCTAC 1696
QY 1054 TCTGAAATTAATATGGTGGCCATGACGGGGTCTACGGGTTTGTGTAGTATCTTAATPACC 1113
DB 1697 TCTGAAATTAATATGGTGGCCATGACGGGGTCTACGGGTTTGTGTAGTATCTTAATPACC 1756
QY 1114 TTAACCCCACTGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATPACCATA 1173
DB 1757 TTAACCCCACTGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATPACCATA 1816
```

```
QY 1174 AGAAT 1178
DB 1817 AGAAT 1821

RESULT 13
ABK83332
ID ABK83332 standard; cDNA; 4685 BP.
XX AC ABK83332;
XX DT 12-AUG-2002 (first entry)
XX DE cDNA encoding human DPRP-1 splice variant #8.
XX KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX OS Homo sapiens.
XX PN WO200231134-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US031874.
XX PR 12-OCT-2000; 2000US-0240117P.
XX PA (FERR ) FERRING BV.
XX PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WI WIPI; 2002-444178/47.
XX DR P-PSDB; ABG61601.
XX PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX PS Disclosure; Page 75-76; 113pp; English.
XX CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders.
XX CC -ABK83343 encode human DPRP proteins
XX SQ Sequence 4685 BP; 1430 A; 853 C; 991 G; 1411 T; 0 U; 0 Other;

Query Match 84.7%; Score 1013.6; DB 6; Length 4685;
Best Local Similarity 91.3%; Pred. No. 3.9e-285;
Matches 1098; Conservative 0; Mismatches 99; Indels 5; Gaps 2;

QY 1 ATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 60
DB 1781 ATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 1840
QY 61 ATCTCTGAGAGGTGACAAAGGCTGACGCGGTGCTACATCTCTGCTGCATCAGTC 120
|||||
```

Db 1841 A<sub>1</sub>CTGAGAGGTGACAGAGCTGACGCGTGGCTACTCACAATCTTGTGTCATCAGTC 1900  
Qy 121 AGCACTGTGACTTCTTTTATAGTAAGTATAGTAAACAGAGAATCCACACTGTGTGCC 180  
Db 1901 AGCACTGTGACTTCTTTTATAGTAAGTATAGTAAACAGAGAATCCACACTGTGTGCC 1960  
Qy 181 TTTTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCA 240  
Db 1961 TTTTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCA 2020  
Qy 241 CCATTTTGGATTTCAG---CAGGTCTCTTCTCTGACT-ATACCTCTCCAGAAAATTTTCTC 295  
Db 2021 CCATTTTGGATTTCAGTCTCAGGTGCGAGTGTGTGAATAATCGTTTAAAGAGTCAAGTA 2080  
Qy 296 TTTTGAAGTACTACTGGAATTTACATTTGTATGGAGTCTCTACAAGCCTCATGATCTACA 355  
Db 2081 TTTCCGGCTTGAATACCTAGCCTCTCTAGTTTATGGTTGTAGTGATAGACACAGGG 2140  
Qy 356 GCTTGGAAAGAAATATCTTACTGTCTGTTCATATATGTTGGTCTCTCAGGTCCTCAATPAGA 415  
Db 2141 ATCTGTGTCACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAATGGGTCAATPAGA 2200  
Qy 416 AATTGACGATCAGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTT 475  
Db 2201 AATTGACGATCAGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTT 2260  
Qy 476 AATCTGTGTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTAAT 535  
Db 2261 AGATCTGTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTAAT 2320  
Qy 536 G<sub>1</sub>AGAGTCAGATATCTTCAAGGTGCTATGCTGGGGCCCAAGTCACTCTGTGGATCTT 595  
Db 2321 GCAGAGTCAGATATCTTCAAGGTGCTATGCTGGGGCCCAAGTCACTCTGTGGATCTT 2380  
Qy 596 CTATGATACAGGATACACGGAACGTTATATGGTCACTGACAGATGACAGGCTA 655  
Db 2381 CTATGATACAGGATACACGGAACGTTATATGGTCACTGACAGATGACAGGCTA 2440  
Qy 656 TTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTACT 715  
Db 2441 TTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTACT 2500  
Qy 716 GCTCTTACATGGTTTCTGGATGAGAAATCTCATTGTCACATACCATGATATTAATCTGAG 775  
Db 2501 GCTCTTACATGGTTTCTGGATGAGAAATCTCATTGTCACATACCATGATATTAATCTGAG 2560  
Qy 776 TTTTCTAGTCAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAG 835  
Db 2561 TTTTCTAGTCAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAG 2620  
Qy 836 CATAAGAGTTCTGGAATCGGAGAACATTAATGAACTGCACTTTTGGCACTACCTTCAAGA 895  
Db 2621 CATAAGAGTTCTGGAATCGGAGAACATTAATGAACTGCACTTTTGGCACTACCTTCAAGA 2680  
Qy 896 AAACCTTTGGATCAGTATGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAATC 955  
Db 2681 AAACCTTTGGATCAGTATGCTGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAATC 2740  
Qy 956 TCTGGTATACATGGCTTATTAACCAAAATGAGAGGTTTAAATCAACAGAAAACACAGAAT 1015  
Db 2741 TCTGGTATACATGGCTTATTAACCAAAATGAGAGGTTTAAATCAACAGAAAACACAGAAT 2800  
Qy 1016 TGATCATCATATTTGATACCTGCCATGTAACATCTACTCTGAAAATFAAATGTGGTGCC 1075  
Db 2801 TGATCATCATATTTGATACCTGCCATGTAACATCTACTCTGAAAATFAAATGTGGTGCC 2860  
Qy 1076 ATGAGGGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAAACCCCATGCTCAAAATC 1135  
Db 2861 ATGAGGGGTCTACGGTTTGTGTAGTAAATCTAATACCTTAAACCCCATGCTCAAAATC 2920  
Qy 1136 AAATGATACATATTTCTGAGAGACCCAGCAATACCAATGAATTTACTAAAAAATAAAAAA 1195  
Db 2921 AAATGATACATATTTCTGAGAGACCCAGCAATACCAATGAATTTACTAAAAAATAAAAAA 2980

Qy 1196 AA 1197  
Db 2981 AA 2982  
RESULT 14  
ACA92421  
ID ACA92421 standard; DNA; 2952 BP.  
XX  
AC ACA92421;  
XX 15-JUL-2003 (first entry)  
XX DNA encoding human PMM-6.  
XX Human; protein modification and maintenance molecule; PMM; cancer;  
KW cell proliferation disorder; atherosclerosis; neurological disorder;  
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;  
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;  
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;  
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnetary;  
KW antiinflammatory; thyromimetic; gene; ds.  
XX Homo sapiens.  
OS  
XX WO2003031939-A2.  
PN  
XX 17-APR-2003.  
PD  
XX 11-OCT-2002; 2002WO-US032850.  
PF  
XX 12-OCT-2001; 2001US-0329689P.  
PR  
XX 25-OCT-2001; 2001US-0335703P.  
PR  
XX 09-NOV-2001; 2001US-0348887P.  
PR  
XX 28-NOV-2001; 2001US-0334145P.  
PR  
XX 06-DEC-2001; 2001US-0337451P.  
PR  
XX 14-DEC-2001; 2001US-0340584P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JK;  
PI Sprague WP, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;  
PI Marquis JP, Yao MG, Richardson TW, Tang TV, Jin P, Chien D;  
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
XX  
XX WPI; 2003-430274/40.  
DR  
XX P-PSDB; ABU92026.  
XX  
XX New human protein modification and maintenance molecules (PMM), useful  
PT for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or  
PT infections.  
PS  
XX Claim 5; Page 285-286; 311pp; English.  
XX  
XX The present invention relates to the isolation of human protein  
CC modification and maintenance molecules (PMM), and the polynucleotide  
CC sequences encoding them. A total of 40 PMM polypeptides (designated PMM  
CC -1 to PMM-40) are disclosed. The sequences of the invention are useful  
CC for diagnosing a condition or disease associated with the expression of  
CC PMM in a subject, preparing a polyclonal or monoclonal antibody, and  
CC generating an expression profile of a sample containing the  
CC polynucleotides. The diseases or conditions associated with decreased  
CC expression or overexpression of PMM are cell proliferation disorders  
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
CC Huntington's disease, stroke), immune/inflammatory disorders (e.g. AIDS,  
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
CC syndrome), gastrointestinal or epithelial disorders, and infections. The  
CC PMM polypeptides or their fragments are useful in screening compounds  
CC for effectiveness as agonists or antagonists of the polypeptides, or in

CC	altering the expression of the target polynucleotide and compounds that	
CC	specifically bind to, or modulate the activity of the polypeptide.	
CC	ACA92416-ACA92455 encode the human PMM polypeptides of the invention	
XX		
SO	Sequence 2952 BP; 886 A; 604 C; 664 G; 798 T; 0 U; 0 Other;	
	Query Match 84.0%; Score 1005.6; DB 10; Length 2952;	
	Best Local Similarity 91.6%; Pred. No. 6.8e-283;	
	Matches 1088; Conservative 0; Mismatches 94; Indels 6; Gaps 2;	
Qy	1 ATTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGTAGCTAGCTAGCTAGCTAA 60	
Db	1771 ATTTGAAGGCACCAAGAGCTCCCTTTAGAGCATCACCTGTAGCTAGCTAGCTAGCTAA 1830	
Qy	61 ATCTGGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTACATTTCTTGCTGCATCAGTC 120	
Db	1831 ATCTGGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTACATTTCTTGCTGCATCAGTC 1890	
Qy	121 AGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180	
Db	1891 AGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 1950	
Qy	181 TTTACAGCTATCAAGTCTGAGATGACCCAACTTGCARAAAGGAAATTTTGGGCCA 240	
Db	1951 TTTACAGCTATCAAGTCTGAGATGACCCAACTTGCARAAAGGAAATTTTGGGCCA 2010	
Qy	241 CCAATTTGGATTCAGCAGGCTCTCTCTGACTATACCTCCGAGAAATTTCTCTTTTG 300	
Db	2011 CCAATTTGGATTCAGCAGGCTCTCTCTGACTATACCTCCGAGAAATTTCTCTTTTG 2070	
Qy	301 AAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCTG 360	
Db	2071 AAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCTG 2130	
Qy	361 GAAAGAAATATCTACTGTCTGTTTATATATGTTGTCCTCAGGGTCAAATAGAAATG 420	
Db	2131 GAAAGAAATATCTACTGTCTGTTTATATATGTTGTCCTCAGGGTCAAATAGAAATG 2187	
Qy	421 ACAGTACAGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATC 480	
Db	2188 ATAATCGTTTAAAGAGGCAAGATTTTCCGCT--TGAATACCCCTAGCCTCTCTAGGTT 2244	
Qy	481 GTGTGGGCATCCACGGCTGCTCTATGAGGAGTACCTCTCCCTGATGGCATTAATGAGA 540	
Db	2245 ATGTGTTGTAGTGATAGACAAACAGGGGATCTGTACCGAGGGCTTAAATTTGAAGGCG 2304	
Qy	541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATG 600	
Db	2305 CCTTAAATATAAATGGTTGCTATTGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATG 2364	
Qy	601 ATACAGGATACCGGAACCTTATATGGGTCACCTGACCAAGTGAACAGGGCTATTACT 660	
Db	2365 ATACAGGATACCGGAACCTTATATGGGTCACCTGACCAAGTGAACAGGGCTATTACT 2424	
Qy	661 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCT 720	
Db	2425 TAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCT 2484	
Qy	721 TACATGGTTCTCGATGAGATGTCATTTGACATACAGTATATTAATGAGTTT 780	
Db	2485 TACATGGTTCTCGATGAGATGTCATTTGACATACAGTATATTAATGAGTTT 2544	
Qy	781 TAGTGGGCTGAAAGCCATATGATTTACAGATCTATCTTCAGAGAGACACAGCATAA 840	
Db	2545 TAGTGGGCTGAAAGCCATATGATTTACAGATCTATCTTCAGAGAGACACAGCATAA 2604	
Qy	841 GAGTTCTGAATCGGAGAACATTTATGAATGCTATCTTTTGCATACCTTCAAGAAACC 900	
Db	2605 GAGTTCTGAATCGGAGAACATTTATGAATGCTATCTTTTGCATACCTTCAAGAAACC 2664	
Qy	901 TTGGATCACGTATTGCTCTTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGG 960	
Db	2665 TTGGATCACGTATTGCTCTTAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGG 2724	

---

Qy	961 TATACACTGGCTATTATAACCAAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTCATC 1020	
Db	2725 TATACACTGGCTATTATAACCAAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATTCATC 2784	
Qy	1021 ATCACATTTTGATACCTGCTGATCAATCATCTACTCTCTGAAATAAATGTGTGCTCATGCA 1080	
Db	2785 ATCACATTTTGATACCTGCTGATCAATCATCTACTCTCTGAAATAAATGTGTGCTCATGCA 2844	
Qy	1081 GGGGTCTACGTTTGTGGTAGTAATCTAATACCTTTAACCCCATGCTTCAAAATCAAATG 1140	
Db	2845 GGGGTCTACGTTTGTGGTAGTAATCTAATACCTTTAACCCCATGCTTCAAAATCAAATG 2904	
Qy	1141 ATACATATTCCTGAGAGACCCAGCAATACCATAGAATTTACTTAAAAA 1188	
Db	2905 ATACATATTCCTGAGAGACCCAGCAATACCATAGAATTTACTTAAAAA 2952	
	RESULT 15	
	AAH15009	
ID	AAH15009 standard; cDNA; 2161 BP.	
XX		
AC	AAH15009;	
XX		
DT	26-JUN-2001 (first entry)	
XX		
DE	Human cDNA sequence SEQ ID NO:12963.	
XX		
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	EP1074617-A2.	
XX		
PD	07-FEB-2001.	
XX		
PF	28-JUL-2000; 2000EP-00116126.	
XX		
PR	29-JUL-1999; 99JP-00248036.	
PR	27-AUG-1999; 99JP-00300253.	
PR	11-JAN-2000; 2000JP-00118776.	
PR	02-MAY-2000; 2000JP-00183767.	
PR	09-JUN-2000; 2000JP-00241899.	
XX		
PA	(HELI-) HELIX RES INST.	
XX		
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;	
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
XX		
DR	WPI; 2001-318749/34.	
XX		
PT	Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.	
PT		
PT		
PT		
PS	Claim 8; SEQ ID NO 12963; 2537pp + Sequence Listing; English.	
XX		
CC	The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer, and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the	

CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention  
XX

SQ Sequence 2161 BP; 662 A; 429 C; 452 G; 618 T; 0 U; 0 Other;

Query Match 83.5%; Score 999.6; DB 4; Length 2161;  
Best Local Similarity 91.5%; Pred. No. 3.4e-281;  
Matches 1082; Conservative 0; Mismatches 94; Indels 6; Gaps 2;

Qy	1	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAA	60
Db	986	ATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAA	1045
Qy	61	ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAATCTTGTGCAATCAGTC	120
Db	1046	ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAATCTTGTGCAATCAGTC	1105
Qy	121	AGCACTGTGACTCTTTTAAAGTAAAGTATAGTAACAGAGAAATCCACACTGTGTCC	180
Db	1106	AGCACTGTGACTCTTTTAAAGTAAAGTATAGTAACAGAGAAATCCACACTGTGTCC	1165
Qy	181	TTTACAAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCCA	240
Db	1166	TTTACAAAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCCA	1225
Qy	241	CCATTTTGGATTGACAGAGTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTG	300
Db	1226	CCATTTTGGATTGACAGAGTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTG	1285
Qy	301	AAAGTACTACTGATTTTACATTTGATGGAGTCTCTACAAGCTCATGATCTACAGCTG	360
Db	1286	AAAGTACTACTGATTTTACATTTGATGGAGTCTCTACAAGCTCATGATCTACAGCTG	1345
Qy	361	GAAGAAATATCTCTGCTGCTTCAATATATGTTGGTGGTCTCTCAGGGTCAATAGAAATG	420
Db	1346	GAAGAAATATCTCTGCTGCTTCAATATATGTTGGTGGTCTCTCAGGGTCAATAGAAATG	1402
Qy	421	AGCATCAGGTGGAGGATCCAAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATC	480
Db	1403	ATAATCGGTTTAAAGGAGTCAAGTATTTCCGCT--TGAATACCCTAGCCTCTCTAGTT	1459
Qy	481	GTGTGGCATCCAGGCTGGTCCATGAGAGTACTCTCCCTGATGGCATTATGCAGA	540
Db	1460	ATGTGGTTGTAGTATAGACAACAGGGGATCCTGTCCAGGGCTTAAATTTGAAGGCG	1519
Qy	541	GGTCAGATATCTTCAGGGTTGCTATTGTGGGGCCCACTGCTGTGGATCTTCTATG	600
Db	1520	CTTTTAAATATAAATGGTTGCTATTGTGGGGCCCACTGCTGTGGATCTTCTATG	1579
Qy	601	ATACAGGATACAGGAACGTTATATGGGTCACTGACAGAAATGAACAGGGCTATTACT	660
Db	1580	ATACAGGATACAGGAACGTTATATGGGTCACTGACAGAAATGAACAGGGCTATTACT	1639
Qy	661	TAGGATCTGTGGCCATGCAAGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT	720
Db	1640	TAGGATCTGTGGCCATGCAAGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCT	1699
Qy	721	TACATGTTTCTCGGATGAGATGTCATTTTGCACATACCAGTATATTACTGAGTTTTT	780
Db	1700	TACATGTTTCTCGGATGAGATGTCATTTTGCACATACCAGTATATTACTGAGTTTTT	1759
Qy	781	TAGTGGGGCTGGAAGGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAA	840
Db	1760	TAGTGGGGCTGGAAGGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAA	1819
Qy	841	GAGTTCCTGAATCGGAGAACATTATGAACTGCTCTTTTGGACTACCTTCAGAAACC	900
Db	1820	GAGTTCCTGAATCGGAGAACATTATGAACTGCTCTTTTGGACTACCTTCAGAAACC	1879

Search completed: May 4, 2006, 09:09:37  
Job time : 686.127 secs

Qy	901	TTGGATCACGTATTGCTGCTCTAAAGTGAATATAATTTTGACCTGTGTAGAACTCTCTGG	960
Db	1880	TTGGATCACGTATTGCTGCTCTAAAGTGAATATAATTTTGACCTGTGTAGAACTCTCTGG	1939
Qy	961	TATACACTGGCTATTTTAAACCAATGAGGAGGTTTAATCAACAGAAACACAGAAATTGATC	1020
Db	1940	TATACACTGGCTATTTTAAACCAATGAGGAGGTTTAATCAACAGAAACACAGAAATTGATC	1999
Qy	1021	ATCACATTTTGTATACCTGCGCATGTAAACATCTACTCTGAAAAATAAATGTTGGCCATGCA	1080
Db	2000	ATCACATTTTGTATACCTGCGCATGTAAACATCTACTCTGAAAAATAAATGTTGGCCATGCA	2059
Qy	1081	GGGGTCTTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAAATCAAATG	1140
Db	2060	GGGGTCTTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAAATCAAATG	2119
Qy	1141	ATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTACT	1182
Db	2120	ATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTACT	2161

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: May 4, 2006, 02:59:51 ; Search time 4483.98 Seconds  
(without alignments)  
12489.839 Million cell updates/sec

Title: US-10-825-632-4  
Perfect score: 1197  
Sequence: 1 atttgaaggcaccacaaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: IDENTITY NUC  
Gap 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_hcc.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_est7.\*  
9: gb\_gss1.\*  
10: gb\_gss2.\*  
11: gb\_gss3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	929.6	77.7	1292 4 AF175225 Homo sapi
2	900	75.2	1265 4 AF176779 Homo sapi
3	832	69.5	1278 4 AF173382 Homo sapi
4	788	65.8	2292 4 CR609512 full-leng
5	768	64.2	864 1 AL542617 AL542617
6	766.2	64.0	804 5 BU687687 UI-CF-EC1
7	758.4	63.4	2649 10 AY411615
8	726.4	60.7	760 1 AL917735
9	717.4	59.9	2649 10 AY411616
10	704.8	58.9	5517 4 AK029788
11	684.8	57.2	3143 4 AK016546
12	677	56.6	735 1 AL043338
13	674.2	56.3	910 2 BG479035
14	673.8	56.3	708 5 BU622228
15	668.8	55.9	673 1 AW303607
16	642	53.6	722 5 BU631054
17	633	52.9	634 1 AI819365
18	625.4	52.2	627 2 BG752328
19	599.2	50.1	1002 2 BE733691
20	598.8	50.0	855 2 BI084885
21	591.4	49.4	597 5 BU784521
22	586.4	49.0	2634 10 AY411617

C 23	578	48.3	639	1	AW958665	EST370735
C 24	577.2	48.2	1042	3	BM557438	AGENCOURT
C 25	574.4	48.0	577	1	AI243857	qh89f09.x
C 26	556.6	46.5	738	7	CN427234	170006000
C 27	551.4	46.1	590	3	BI711119	id90q08.x
C 28	549	45.9	804	7	CN438529	BE0401681
C 29	533.4	44.6	776	6	CB522171	UI-M-GHO-
C 30	532.4	44.5	905	6	CB209486	AGENCOURT
C 31	529.4	44.2	536	3	BI710784	id90g08.y
C 32	519.2	43.4	719	8	DR768184	ILLUMIGEN
C 33	509.8	42.6	746	5	BY764370	BY764370
C 34	502.8	42.0	800	2	BG921267	602824107
C 35	496.8	41.5	500	1	AA496257	aa24b08.s
C 36	493.6	41.2	520	1	AW469326	hc81h10.x
C 37	490	40.9	617	2	BE789604	601481552
C 38	483.6	40.4	970	2	BG403555	602419260
C 39	483	40.4	494	1	AA278626	zs78g11.s
C 40	482.2	40.3	797	6	CA749361	UI-M-FY0-
C 41	482	40.3	493	1	AA278625	zs78g11.f
C 42	476.2	39.8	1038	1	AL582206	AL582206
C 43	474.4	39.6	820	6	CF532140	UI-M-GHO-
C 44	473.2	39.5	554	1	AW469635	hd32b10.x
C 45	468.2	39.1	835	7	CV557345	UI-M-H20-

ALIGNMENTS

RESULT 1  
AF175225 1292 bp mRNA linear HTC 01-AUG-2003  
LOCUS AF175225 Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.  
DEFINITION AF175225  
ACCESSION AF175225  
VERSION AF175225.1 GI:33338055  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1292)  
AUTHORS Zhao, B., Xu, H.S., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B., Wang, X.Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J., Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zhang, W.Y., Teng, C.Y., Liu, Q., Yu, L.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hui, R.T.  
TITLE Direct Submission  
JOURNAL Submitted (04-AUG-1999) Molecular Medicine Center for Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167, Bei Li Shi Lu, Beijing 100037, P.R. China  
FEATURES  
source location/Qualifiers  
1..1292  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="aorta"  
142..1023  
/codon\_start=1  
/product="MSTP135"  
/protein\_id="AA013650.1"  
/db\_xref="GI:33338055"  
/translation="MTQLAKQRNFPPWQQVFLTLLQKFSLLKVLKLLDLCMGCS  
TSPhgSTGLERNILLLFTYGGKVLQVNNRFGKFFRLNTLPGLVYVVGIDNRGS  
CPRLKPEGAFFKMGQIEIDQVGLQYLAQRDFIDLRVGHGVSYGYSGLMAL  
MORSIDIFRVAIAGAPVTWTFDTGTERYMGHPDQEQYVLGSVMAQKFPSEPN  
RLLLHGFLENVFAHTSILLSFLVRACKPYDLQIYQERHSIRVPESGHEYLHL  
HYLQENLGSRIAAKVI"

ORIGIN  
Query Match 77.7%; Score 929.6; DB 4; Length 1292;  
Best Local Similarity 87.1%; Pred. No. 4e-246;  
Matches 1119; Conservative 0; Mismatches 15; Indels 15; Gaps 3;  
64 CTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCATTCTTGTGTCATGCTAGC 123

Db 1 CTGAGAGGTGACAGGCTGACTGACCGTGGGTACTCACAATCTTGCTGCATCAGTCAGC 60  
Qy 124 ACTGTGACTCTTTTAAAGTAAGTATAGTAACACAGAGAATCCACACTGTGTGTCCTTTT 183  
Db 61 ACTGTGACTCTTTTAAAGTAAGTATAGTAACACAGAGAATCCACACTGTGTGTCCTTTT 120  
Qy 184 ACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAACAAAGGAATTTTGGGCCACCA 243  
Db 121 ACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAACAAAGGAATTTTGGGCCACCA 180  
Qy 244 TTTTGGATTTCAGCAGGTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAA 303  
Db 181 TTTTGGATTTCAGCAGGTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAA 240  
Qy 304 GTACTACTGATTATCAATTTGATGGGATGCTCTCAAGCCTCATGATCTACA---GCCTG 360  
Db 241 GTACTACTGATTATCAATTTGATGGGATGCTCTCAAGCCTCATGATCTACAGGCCG 300  
Qy 361 GAAAGAAATATCTACTGTGCTGTTCATATATGTTGCTCTCA--- 403  
Db 301 GAAAGAAATATCTACTGTGCTGTTCATATATGTTGCTCTCA--- 360  
Qy 404 ----- 403  
Db 361 AATCGGTTTAAAGGAGTCAAGTTTTCGCTTGAATACCTTACCTTTTAGTTATGTG 420  
Qy 404 ----- 403  
Db 421 GTTGNAGGATAGACAACAGGGATCTGTCCCGAGGGCTTAATTTGAAGGCCCTTT 480  
Qy 404 -----GGTCAATATAGAAATGACGATCAGGTGGAAGGACTCCATATCTAGCT 452  
Db 481 AAATTTAAATGGGTCAATATAGAAATGACGATCAGGTGGAAGGACTCCATATCTAGCT 540  
Qy 453 TCTCGATATGATTTCATTTAGATCTGTGCGGATCCACGCTGCTGTCTATGAGGGA 512  
Db 541 TCTCGATATGATTTCATTTAGATCTGTGCGGATCCACGCTGCTGTCTATGAGGGA 600  
Qy 513 TACCTCTCCCTGATGCTTAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGG 572  
Db 601 TACCTCTCCCTGATGCTTAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGG 660  
Qy 573 GCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTGTTATGAGGTCAC 632  
Db 661 GCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTGTTATGAGGTCAC 720  
Qy 633 CCTGACAGAAATGAACAGGCTATTACTTATGATCTGTGGCCATGCAAGCAGAAAGTTC 692  
Db 721 CCTGACAGAAATGAACAGGCTATTACTTATGATCTGTGGCCATGCAAGCAGAAAGTTC 780  
Qy 693 CCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGGATGGAATGTCCATTTT 752  
Db 781 CCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGGATGGAATGTCCATTTT 840  
Qy 753 GCATATACCAATATTTACTGAGTCTTTTACTGAGGCTGGAAGCCATATGATTTACAG 812  
Db 841 GCATATACCAATATTTACTGAGTCTTTTACTGAGGCTGGAAGCCATATGATTTACAG 900  
Qy 813 ATCTATCTCAGGAGACACAGCATTAAGTCTGATCGGAGACATATGACTG 872  
Db 901 ATCTATCTCAGGAGACACAGCATTAAGTCTTCTGATCGGAGACATATGACTG 960  
Qy 873 CATCTTTGCACTACCTTCAAGAAAACCTTGATACAGTATGCTGCTCTTAAAGTGATA 932  
Db 961 CATCTTTGCACTACCTTCAAGAAAACCTTGATACAGTATGCTGCTCTTAAAGTGATA 1020  
Qy 933 TAAATTTGACTGTGTAGAACTCTCTGGTATACACTGGGCTATTAAACAAATGAGGAGT 992  
Db 1021 TAAATTTGACTGTGTAGAACTCTCTGGTATACACTGGGCTATTAAACAAATGAGGAGT 1080  
Qy 993 TTAATCAACAGAAAACACAGAAATTTGATCTATCAATTTTGATACCTGCGCATCTACTA 1052

Db 1081 TTAATCAACAGAAAACACAGAAATTTGATCTACATTTTGATACCTGCCATGTAACATCTA 1140  
Qy 1053 CTCTGTAATAATAATGTTGGTCCATGTCAGGGGTCTAGGTTTGTAGTAATCTAATAC 1112  
Db 1141 CTCTGTAATAATAATGTTGGTCCATGTCAGGGGTCTAGGTTTGTAGTAATCTAATAC 1200  
Qy 1113 CTTAACCCCATGCTCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCATACAT 1172  
Db 1201 CTTAACCCCATGCTCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCATACAT 1260  
Qy 1173 AAGAATTTACTTAAAAAATTTTTTTTTT 1197  
Db 1261 AAGAATTTCTTAAAAAATTTTTTTTTT 1285  
RESULT 2  
AF176779  
LOCUS Homo sapiens MSTP141 mRNA, complete cds. linear HTC 01-AUG-2003  
DEFINITION  
AF176779  
ACCESSION AF176779.1 GI:33338069  
VERSION  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 1265)  
Hui, R. T., Liu, Y. Q., Wang, X. Y., Qin, B. M. and Sheng, H.  
TITLE Homo sapiens normal aorta mRNA MST141  
JOURNAL Unpublished  
REFERENCE  
2 (bases 1 to 1265)  
Hui, R. T., Liu, Y. Q., Wang, X. Y., Qin, B. M. and Sheng, H.  
TITLE Direct Submission (10-AUG-1999) Molecular Medicine Center for  
JOURNAL Submitted (10-AUG-1999) Cardiovascular Institute, CAMS & PUMC, 167,  
Cardiovascular Disease, Beijing, 100037, P.R. China  
Be Li Shi Lu, Beijing, 100037, P.R. China  
FEATURES  
source  
1. .1265  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="aorta"  
268..741  
/codon\_start=1  
/product="MSTP141"  
/protein\_id="AAQ13657.1"  
/db\_xref="GI:33338070"  
/translation="MLYKPHDLQPGKKYPTVLFIYGGPQVQVNNRFKGVKYPRLNTL  
ASLGVVVVVIDNRGSGHRGLKFEKGYKMGQIBIDDOEGLOYLAGRYDFIDLDVRG  
IHGWSYGVVLSLMLMQRSDIFRVAIAGAPVTLWIFDYDTGYTRYMGHPDQNSQG"  
ORIGIN  
Query Match 75.2%; Score 900; DB 4; Length 1265;  
Best Local Similarity 86.2%; Pred. No. 6.5e-238;  
Matches 1105; Conservative 0; Mismatches 5; Indels 172; Gaps 3;  
Qy 64 CTGAGAGGTGACAGG-CTGACTGACCGTGGTACTCACAATCTTGCTGCATCAGTCAG 122  
Db 1 CTGAGAGGTGACAGGCTGACTGACCGTGGTACTCACAATCTTGCTGCATCAGTCAG 60  
Qy 123 CACTGTGACTCTTTTATAAGTAAGTATAGTAACACAGAGAATCCACACTGTGTGCTCTT 182  
Db 61 CACTGTGACTCTTTTATAAGTAAGTATAGTAACACAGAGAATCCACACTGTGTGCTCTT 120  
Qy 183 TAGAAGTATCAAGTCTGAGATGACCCCACTTCGAAACAAAGGAATTTTGGGCCACC 242  
Db 121 TAGAAGTATCAAGTCTGAGATGACCCCACTTCGAAACAAAGGAATTTTGGGCCACC 180  
Qy 243 ATTTTGAATTCAGAGGTCTCTCTGACTACTCTCTGAGAAATTTTCTCTTTGAA 302  
Db 181 ATTTTGAATTCAGAGGTCTCTCTGACTACTCTCTGAGAAATTTTCTCTTTGAA 240



```
Db 301 AGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGTGCAGTTGGTGAATATTC 360
Qy 404 ----- 403
Db 361 GGTTTAAAGGAGTCAAGTATTTTCGGCTTTGAATACCTAGCCTCTCTAGGTATGTGGTTG 420
Qy 404 ----- 403
Db 421 TAGTGATAGACACAGGGGATCTGTGTACCAGGGCTTAAATTTGAAGGCGCTTTAAAT 480
Qy 404 -----GGGTCAAATAGAAATGACGATCAGGTGGAAGACTCCCAATATCTAGCTTTC 456
Db 481 ATAAATGGGTCAAATAGAAATGACGATCAGGTGGAAGACTCCCAATATCTAGCTTTC 540
Qy 457 GATATG-ATTTCATTGACTAGATCGTGTGGGCATCCACGCTGGTCTCTATGAGGATAC 515
Db 541 GATATGAATTTTCATTGACTAGATCGTGTGGGCATTTCCCGCTGGTCTCTATGAGATAC 600
Qy 516 CTCTCCCTGATGGCATTATGACAGAGGTACAGATATCTTCAGGGTGTCTATGCTGGGGC 575
Db 601 CCTCTCCTGATGGCAT--ATGCAGAGGTACAGATATCTTCAGG--TGTATTGCTGG--C 654
Qy 576 CCAGTCACCTCTGGATCTTCTATGATACAGATACAGGAAAGTATATGTTGGTCACTC 635
Db 655 CCGTCCACCTCTGTGATCTCTAT--GAACAGATACAGG-ACGTATATGTTGGTCACTG 710
Qy 636 GACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCMAGCAGAAAGTTCCCC 695
Db 711 ACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCC 770
Qy 696 TCTGAACCAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAAATGTCATTGTGCA 755
Db 771 TTTGAACCAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAAATGTCATTGTGCA 830
Qy 756 CATACCAATATATTACTGAGTTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATC 815
Db 831 CATACCAATATATTACTGAGTTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATC 890
Qy 816 TATCCTCAGGAGAGACAGCATAGAGTTCTCTGAATCGGGAGAACATTATGAACTGCAT 875
Db 891 TATCCTCAGGAGAGACAGCATAGAGTTCTCTGAATCGGGAGAACATTATGAACTGCAT 950
Qy 876 CTTTTCACATCTCTCAAGAAAACCTTGGATCAGTATTTGCTCTTAAAGAGTATATAA 935
Db 951 CTTTTCACATCTCTCAAGAAAACCTTGGATCAGTATTTGCTCTTAAAGAGTATATAA 1010
Qy 936 TTTTGACCTGTGAGAACTCTCTGTATACACTGGCTATTAAACCAATGAGGAGTTTA 995
Db 1011 TTTTGACCTGTGAGAACTCTCTGTATACACTGGCTATTAAACCAATGAGGAGTTTA 1070
Qy 996 ATCAACAGAAACACAGAAATGATCATCAGATTTTGTATACCTGCCATGTAACATCTACTC 1055
Db 1071 ATCAACAGAAACACAGAAATGATCATCAGATTTTGTATACCTGCCATGTAACATCTACTC 1130
Qy 1056 CTGAAATAAATGTGGTGCATGACGGGCTCTACGGTTTGTGGTAGTAATCTTAATACCTT 1115
Db 1131 CTGAAATAAATGTGGTGCATGACGGGCTCTACGGTTTGTGGTAGTAATCTTAATACCTT 1190
Qy 1116 AACCCACATGCTCAAAATCAAATGATACATATTTCTTGAGAGCCACCAATACCAATAG 1175
Db 1191 AACCCACATGCTCAAAATCAAATGATACATATTTCTTGAGAGCCACCAATACCAATAG 1250
Qy 1176 AATTACTAAAAA 1197
Db 1251 AATTACTAAAAA 1272

RESULT 4
CR609512 2292 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DL005YD02 of B cells (Ramos cell line)
DEFINITION Cot 25-normalized of Homo sapiens (human).
ACCESSION CR609512
```

```
VERSION CR609512.1 GI:50490319
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2292)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue
REFERENCE 2 (bases 1 to 2292)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source Location/Qualifiers
1..2292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YD02"
/issue_type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 65.8%; Score 788; DB 4; Length 2292;
Best Local Similarity 85.5%; Pred. No. 8.6e-207;
Matches 973; Conservative 0; Mismatches 0; Indels 165; Gaps 2;
Qy 1 ATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAA 60
Db 1320 ATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTAA 1379
Qy 61 ATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTCTTGTGCATCAGTC 120
Db 1380 ATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTCTTGTGCATCAGTC 1439
Qy 121 AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTC 180
Db 1440 AGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTC 1499
Qy 181 TTTCAGAGCTATCAAGTCTCTGAGATGACCAACTTGCAAAACAAAGAAATTTGGGCCA 240
Db 1500 TTTCAGAGCTATCAAGTCTCTGAGATGACCAACTTGCAAAACAAAGAAATTTGGGCCA 1559
Qy 241 CCATTTTGGATTACAGAGTCTCTTCTGACTATATCTCTCGAGAAATTTCTCTTTTG 300
Db 1560 CCATTTTGGATTACAGAGTCTCTTCTGACTATATCTCTCGAGAAATTTCTCTTTTG 1619
Qy 301 AAAGTACTACTGGATTACATTGTATGGATGCTCTACAGCCTCATGATCTACAGCTG 360
Db 1620 AAAGTACTACTGGATTACATTGTATGGATGCTCTACAGCCTCATGATCTACAGCTG 1679
Qy 361 GAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGGTCAAAATGAAATTG 420
Db 1680 GAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCA----- 1722
Qy 421 ACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATC 480
Db 1723 ----- 1722
Qy 481 GTGTGGGCATCCACGGTGGTCTTATGGAGGATACCTCTCCCTGATGCCATTAATGAGAA 540
```

Db 1723 ----- 1722

Qy 541 GGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCAGTCACCTCTGTGGATCTTCTATG 600

Db 1723 -----GGTTGCTATTGCTGGGGCCAGTCACCTCTGTGGATCTTCTATG 1766

Qy 601 ATACAGGATACACGGAACGTTATATGCGTACCTGACCAAGTAAACAGGGCTATTACT 660

Db 1767 ATACAGGATACACGGAACGTTATATGCGTACCTGACCAAGTAAACAGGGCTATTACT 1826

Qy 661 TAGGATCTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACAAATCGTTTACTGCTCT 720

Db 1827 TAGGATCTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACAAATCGTTTACTGCTCT 1886

Qy 721 TACATGGTTTCTCGGATGAGAAATGTCATTTTGGACATACCAAGTATATTACTGAGTTTTT 780

Db 1887 TACATGGTTTCTCGGATGAGAAATGTCATTTTGGACATACCAAGTATATTACTGAGTTTTT 1946

Qy 781 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTTCAGGAGAGACACAGCATAA 840

Db 1947 TAGTGAGGCTGGAAGCCATATGATTTA-----CAGGAGAGACACAGCATAA 1994

Qy 841 GAGTTCTGAACTGGGAGACATATGAACTGCACTTTTGCACCTCTCAAGAAACC 900

Db 1995 GAGTTCTGAACTGGGAGACATATGAACTGCACTTTTGCACCTCTCAAGAAACC 2054

Qy 901 TTGGATCACGATTTGCTGCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG 960

Db 2055 TTGGATCACGATTTGCTGCTTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGG 2114

Qy 961 TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAATTGATC 1020

Db 2115 TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAACACAGAATTGATC 2174

Qy 1021 ATCACATTTTGATACCTGCATGTAAACATCTACTCTGAAAATAAATGTGGTGCATGCA 1080

Db 2175 ATCACATTTTGATACCTGCATGTAAACATCTACTCTGAAAATAAATGTGGTGCATGCA 2234

Qy 1081 GGGGCTACGGTTTGGGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAA 1138

Db 2235 GGGGCTACGGTTTGGGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAA 2292

RESULT 5

AL542617

LOCUS

AL542617 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE014YN22

DEFINITION

5-PRIME, mRNA sequence.

ACCESSION

AL542617

VERSION

AL542617.3 GI:45718193

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 864)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 15, 2001 this sequence version replaced gi:30547946.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 7542.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DE014DG11Q1P1&c=7542.r.

FEATURES

source

Location/Qualifiers

1..864

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DE014YN22"

/tissue\_type="PLACENTA"

/clone\_lib="Homo sapiens PLACENTA"

/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 64.2%; Score 768; DB 1; Length 864;

Best Local Similarity 98.9%; Pred. No. 2.3e-201;

Matches 787; Conservative 6; Mismatches 1; Indels 2; Gaps 2;

Qy 404 GGGTCAAAATAGAAATTTGAGATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGA 463

Db 64 GGGTCAAAATAGAAATTTGAGATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGA 123

Qy 464 TTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCT 523

Db 124 TTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCT 183

Qy 524 GATGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTGTCTATTGCTGGGGCCAGTCCAC 583

Db 184 GATGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTGTCTATTGCTGGGGCCAGTCCAC 243

Qy 584 TCTGTGGATCTTCTATGATACAGGATACACGAAAGTATATGAGTCACTCTGACAGAA 643

Db 244 TCTGTGGATCTTCTATGATACAGGATACACGAAAGTATATGAGTCACTCTGACAGAA 303

Qy 644 TGAACAGGGCTATCTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACC 703

Db 304 TGAACAGGGCTATCTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACC 363

Qy 704 AAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGACATACCCAG 763

Db 364 AAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGACATACCCAG 423

Qy 764 TATATTACTGAGTTTTTTAGTGAGGGCTGAAAAGCCATATGATTTTACAGATCTATCTCA 823

Db 424 TATATTACTGAGTTTTTTAGTGAGGGCTGAAAAGCCATATGATTTTACAGATCTATCTCA 483

Qy 824 GGAGAGACACAGCATTAAGAGTTCTTGAATCGGGAGAAATTAATGAACTGTCATCTTTTGA 883

Db 484 GGAGAGACACAGCATTAAGAGTTCTTGAATCGGGAGAAATTAATGAACTGTCATCTTTTGA 543

Qy 884 CTACCTTCAAGAAAACCTTGGATCAGTATTTGCTGCTCTAAAAGTGATATAATTTTGACC 943

Db 544 CTACCTTCAAGAAAACCTTGGATCAGTATTTGCTGCTCTAAAAGTGATATAATTTTGACC 603

Qy 944 TGTGTAGAATCTCTCTGTATACACTGGCTATTTTAAACAAATGAGGAGTTTAAATCAACAG 1003

Db 604 TGTGTAGAATCTCTCTGTATACACTGGCTATTTTAAACAAATGAGGAGTTTAAATCAACAG 663

Qy 1004 AAAACACAGAAATTTGATCATCAATTTTGTATCTCTGCCATGTAACTCTACTCTGAAAT 1063

Db 664 AAAACACAGAAATTTGATCATCAATTTTGTATCTCTGCCATGTAACTCTACTCTGAAAT 723

Qy 1064 AAATGCTGGTCCATGACAGGGTCTAGGTTGTTGGTAGTAATCTAATACCTTTAACCCAC 1123

Db 724 AAATGCTGGTCCATGACAGGGTCTAGGTTGTTGGTAGTAATCTAATACCTTTAACCCAC 783

Qy 1124 ATGCTCAAAATCAA-TGATACATATTC-TGAGAGACCCAGCAATACCAATGAATATAC 1181

Db 784 ATGCTCAAAATCAAATTTGATCATATTTCTTGGAGACCCAGCAATACCAATGAATATAC 843

Qy 1182 TAAAAAATCAAAATCA 1197

|||||

```

Db      844 CAAAAAAAAAAAAAAAA 859

RESULT 6
BU687687/c
LOCUS   BU687687
DEFINITION
UT-CF-EC1-adu-o-02-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
UT-CF-EC1-adu-o-02-0-UI 3', mRNA sequence.
ACCESSION
VERSION  BU687687.1
KEYWORDS  BU687687.1 GI:23543781
SOURCE    EST.
ORGANISM  Homo sapiens (human)

REFERENCE
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
           discovery
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
PUBMED     8889548
COMMENT    Contact: McCray, PB
           McCray Lab
           University of Iowa
           2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
           Tel: 319 356 4866
           Fax: 319 356 7171
           Email: paul-mccray@uiowa.edu
           Tissue Procurement: Dr. M. J. Welsh, University of Iowa
           cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
           cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
           DNA sequencing by: Dr. M. Bento Soares, University of Iowa
           Clone Distribution: Researchers may obtain clones from Research
           Genetics (www.regen.com) or from Open Biosystems
           Seq primer: M13 FORWARD
           POLYA=yes.

FEATURES             Location/Qualifiers
     source           1..804
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EC1-adu-o-02-0-UI"
                     /tissue_type="Lung"
                     /dev_stage="Adult and Fetal"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_libs="UI-CF-EC1"
                     /note="Organ: Lung; Vector: p77T3-Pac (Pharmacia) with a
                     modified polylinker; Site 1: EcoR I; Site 2: Not I;
                     UI-CF-EC1 is a normalized cDNA library containing the
                     following tissue(s): Normal lung from adult and from fetal
                     day 64, day 87, week 19 and week 42. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806 1996, First strand cDNA,
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor digested with Not I, and cloned directionally
                     into p77T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     GATC tail. The sequence tag for this library is
                     AAGTGCCTTAC.
                     TAG-TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
                     and 380-383
                     TAG_LiB=UI-CF-EC1
                     TAG_SEQ=AAGTGCCTTAC"

ORIGIN
Query Match      64.0%; Score 766.2; DB 5; Length 804;
Best Local Similarity 99.4%; Pred. No. 7.2e-201;
Matches 790; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

1183 AAAAAAAAAAAAAAAAAA 1197
18  AAAAAAAAAAAAAAAAAA 4

RESULT 7
AY411615
LOCUS   AY411615
DEFINITION
Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION
VERSION  AY411615.1
KEYWORDS  AY411615.1 GI:39767583
SOURCE    GSS.
ORGANISM  Homo sapiens (human)

REFERENCE
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Eumarchontoglires; Primates; Catarrhini;
           Homnidae; Homo.
           1 (bases 1 to 2649)
           Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
           Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

```

```

403 AGGCTCAATAGAAATTGACGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATG 462
796 ATGGTCAATAGAAA-TGACGATCA-GTGAAGGACTCCAATATCTAGCTTCTCGATATG 739
463 ATTTCAITGACTTAGATTCGTGTGGGATCCACGGCTGGTCCATATGGAGGATACCTCTCCC 522
738 ATTTCAITGACTTAGATTCGTGTGGGATCCACGGCTGGTCCATATGGAGGATACCTCTCCC 679
523 TGATGGCAITTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCA 582
678 TGATGGCAITTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCA 619
583 CTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTTATATGGGTCACTCTGACAGA 642
618 CTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTTATATGGGTCACTCTGACAGA 559
643 ATGAACAGGGCTATTACTTTAGGATCTGTGGCCATCAAGCAAGTATATGGGTCACTCTGAC 702
558 ATGAACAGGGCTATTACTTTAGGATCTGTGGCCATCAAGCAAGTATATGGGTCACTCTGAC 499
703 CAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTGACATACCA 762
498 CAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTGACATACCA 439
763 GTATATTACTGAGTTTCTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTC 822
438 GTATATTACTGAGCTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTC 379
823 AGGAGACACACGACATAGAGTTCCTGAATCGGAGAACATATGAATGATCTGCTTTGTC 882
378 AGGAGACACACGACATAGAGTTCCTGAATCGGAGAACATATGAATGATCTGCTTTGTC 319
883 ACTACCTTCAAGAAAAACCTTGGATCAGGTATGCTGCTCTAAAGTGATATAAATTTGAC 942
318 ACTACCTTCAAGAAAAACCTTGGATCAGGTATGCTGCTCTAAAGTGATATAAATTTGAC 259
943 CTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGGTTTAATCAACA 1002
258 CTGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGGTTTAATCAACA 199
1003 GAAAAACACAGAAATTGATCATCATCTTTTGATACCTGCCATGTAACATCTACTCTCGAAAA 1062
198 GAAAAACACAGAAATTGATCATCATCTTTTGATACCTGCCATGTAACATCTACTCTCGAAAA 139
1063 TAAATGTGTGGTCATGCAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCA 1122
138 TAAATGTGTGGTCATGCAGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCA 79
1123 CATGCTCAAAATCAATGATACATATTCCTGAGAGCCAGCAATACCAATGAATTAAT 1182
78 CATGCTCAAAATCAATGATACATATTCCTGAGAGCCAGCAATACCAATGAATTAAT 19
1183 AAAAAAAAAAAAAAAAAA 1197
18  AAAAAAAAAAAAAAAAAA 4

```

Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 Science 302 (5652), 1960-1963 (2003)  
 14671302  
 2 (bases 1 to 2649)  
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source  
 1..2649  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>2649  
 /gene="DPP8"  
 /locus\_tag="HCM4283"  
 ORIGIN  
 Query Match 63.4%; Score 758.4; DB 10; Length 2649;  
 Best Local Similarity 84.7%; Pred. No. 1.4e-198;  
 Matches 916; Conservative 0; Mismatches 19; Indels 147; Gaps 1;  
 QY 1 ATTTTGAAGGCCAACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 60  
 DB 1568 ATTTTGAAGGCCAACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTAA 1627  
 QY 61 ATCTGGAGAGTGCAAGGCTGACTGACCGTGGCTACTCACAATCTTCTGTCATCAGTC 120  
 DB 1628 ATCTGGAGAGTGCAAGGCTGACTGACCGTGGCTACTCACAATCTTCTGTCATCAGTC 1687  
 QY 121 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTC 180  
 DB 1688 AGCACTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTC 1747  
 QY 181 TTACACGCTATCAAGTCTCTGAAGATGACCCCACTTGCAGAAACAAAGAAATTTGGGCCA 240  
 DB 1748 TTACACGCTATCAAGTCTCTGAAGATGACCCCACTTGCAGAAACAAAGAAATTTGGGCCA 1807  
 QY 241 CCAATTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTG 300  
 DB 1808 CCAATTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTG 1867  
 QY 301 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTG 360  
 DB 1868 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTG 1927  
 QY 361 GAAAGAAATATCCTACTGTGCTGTCTATATATGCTGCTCTCA----- 403  
 DB 1928 GAAAGAAATATCCTACTGTGCTGTCTATATGCTGCTCTCAAGTGGTGGTAATA 1987  
 QY 404 ----- 403  
 DB 1988 ATCGGTTTTAAAGGAGTCAAGATTTTCGGCTTGAATACCCTAGCCTCTTAGTATGTGG 2047  
 QY 404 ----- 403  
 DB 2048 TTGTAGTGATAGACAACAGGGATCTGTGCACCGAGGGCTTAAATTTGAAGCGGCTTTTA 2107  
 QY 404 -----GGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453  
 DB 2108 AATATAAAATGGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 2167  
 QY 454 CTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCA CGGTGGTCTTATGAGGAT 513  
 DB 2168 CTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCA CGGTGGTCTTATGAGGAT 2227

QY 514 ACCTCTCCCTGATGGCATTAAATCAGAGGTCAGATATCTTCAGGGTGTCTATTGCTGGGG 573  
 DB 2228 ACCTCTCCCTGATGGCATTAAATCAGAGGTCAGATATCTTCAGGGTGTCTATTGCTGGGG 2287  
 QY 574 CCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAGCAAGCTTATATGGGTCAACC 633  
 DB 2288 CCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAGCAAGCTTATATGGGTCAACC 2347  
 QY 634 CTGACCAAGATGAACACGGCTATTACTTAGGATCTCTGGCCATGCAAGCAGCAAAAAATTTCC 693  
 DB 2348 CTGACCAAGATGAACACGGCTATTACTTAGGATCTCTGGCCATGCAAGCAGCAAAAAATTTCC 2407  
 QY 694 CCTCTGAACCAAAATCGTTTACTCTCTTACATGGTTTCTCTGATGAGAAATGTCATTTTG 753  
 DB 2408 NNNNNNACCAAAATCGTTTACTCTCTTACATGGTTTCTCTGATGAGAAATGTCATTTTG 2467  
 QY 754 CACATACCAAGTATATTAATCTGAGTTTTTATGAGGGCTGGAAGCCATATGATTTTACAGA 813  
 DB 2468 CACATACCAAGTATATTAATCTGAGTTTTTATGAGGGCTGGAAGCCATATGATTTTACAGA 2527  
 QY 814 TCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATATATGAACATGC 873  
 DB 2528 TCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCTCGAATCTCGAAGAACATATATGAACATGC 2587  
 QY 874 ATCTTTTGCACCTACCTTTCAAGAAAAACCTTGGATCAGTATTTCTCTCTAAAAGTGATAT 933  
 DB 2588 ATCTTTTGCACCTACCTTTCAAGAAAAACCTTGGATCAGTATTTCTCTCTAAAAGTGATAT 2647  
 QY 934 AA 935  
 DB 2648 AA 2649  
 RESULT 8  
 AI917735/c  
 LOCUS  
 DEFINITION  
 t11d07.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2240461 3'  
 similar to TR:075273 075273 R26984\_1 ;, mRNA sequence.  
 ACCESSION  
 AI917735  
 VERSION  
 AI917735.1 GI:5637590  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 760)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 Tumor Gene Index  
 JOURNAL  
 Unpublished (1997)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
 R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html  
 Insert Length: 1041 Std Error: 0.00  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 444.  
 location/Qualifiers  
 source  
 . . . . .1..760  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2240461"  
 /tissue\_type="pooled germ cell tumors"



/lab\_host="DH10B"  
/clone\_lib="NCI CGAP GC6"  
/notes="Vector: pRT3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA  
from the normalized library NCI CGAP GC4 was prepared, and  
ss circles were made in vitro. Following HAP purification,  
this DNA was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from a pool  
of 5,000 clones made from the same library (clones IDs  
1257096-1258631, 1469064-1470983, and 1475592-1476743).  
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 60.7%; Score 726.4; DB 1; Length 760;  
Best Local Similarity 96.8%; Pred. No. 7, 7e-190;  
Matches 736; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
QY 429 GTGGAGGAGCTCAATATCTAGCTCTCGATGATGATTCATGCTAGATCTGTGGGC 488  
DB 760 GTGGAGGAGCTCAATATCTAGCTCTCGATGATGATTCATGCTAGATCTGTGGGC 701  
QY 489 ATCCACGGCTGCTCTATGGAGATACCTCTCCGATGGCATTAATGCGAGGTGAGT 548  
DB 700 ATCCACGGCTGCTCTATGGAGATACCTCTCCGATGGCATTAATGCGAGGTGAGT 641  
QY 549 ATCTTCAGGTTGCTATGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACGGA 608  
DB 640 ATCTTCAGGTTGCTATGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACGGA 581  
QY 609 TACACGGAACGTTATATGCTGGGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCT 668  
DB 580 TACACGGAACGTTATATGCTGGGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCT 521  
QY 669 GTGGCCATGAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGT 728  
DB 520 GTGGCCATGAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGT 461  
QY 729 TTCTGTCATGAGAAATGTCCATTTTGCATACACAGTATATTACTGAGTTTTTGTAGGAG 788  
DB 460 TTCTGTCATGAGAAATGTCCATTTTGCATACACAGTATATTACTGAGTTTTTGTAGGAG 401  
QY 789 GCTGGAAGCCATATGATTTACAGATCTATCTCTGAGGAGACACAGACATAAGAGTTTCCT 848  
DB 400 GCTGGAAGCCATATGATTTACAGATCTATCTCTGAGGAGACACAGACATAAGAGTTTCCT 341  
QY 849 GAATCGGGAGAACATTTATGATGATCTTTTGCATCTTCTGCACTACCTTCAAGAAACCTTGGATCA 908  
DB 340 GAATCGGGAGAACATTTATGATGATCTTTTGCATCTTCTGCACTACCTTCAAGAAACCTTGGATCA 281  
QY 909 CGTATTCGCTCTTAAAGTGATATAATTTTGCACCTGTGTAGTACTCTCTGTTATACAT 968  
DB 280 CGTATTCGCTCTTAAAGTGATATAATTTTGCACCTGTGTAGTACTCTCTGTTATACAT 221  
QY 969 GCCTATTTAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATGATCATCACTT 1028  
DB 220 GCCTATTTAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATGATCATCACTT 161  
QY 1029 TTGATACCTGCGATTAACATCTATCTCTGAAATAAATATGTTGGTGCCATGCGGGGTCTA 1088  
DB 160 TTGATACCTGCGATTAACATCTATCTCTGAAATAAATATGTTGGTGCCATGCGGGGTCTA 101  
QY 1089 CGGTTTGTGGTAGTAATCTTAATACCTTTAAACCCACATGCTCAAAATCAATGATACATAT 1148  
DB 100 CGGTTTGTGGTAGTAATCTTAATACCTTTAAACCCACATGCTCAAAATCAATGATACATAT 41  
QY 1149 TCCTGAGAGACCCAGCAATACCATTAAGAATTACTAAAAA 1188  
DB 40 TCCTGAGAGACCCAGCAATACCATTAAGAATTACTAAAAA 1

RESULT 9  
AY411616  
LOCUS

AY411616 2649 bp DNA linear GSS 16-DEC-2003

DEFINITION Pan troglodytes DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY411616  
VERSION AY411616.1 GI:39767584  
KEYWORDS GSS.  
SOURCE Pan troglodytes (Chimpanzee)  
ORGANISM Pan troglodytes  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Pan.  
REFERENCE 1 (bases 1 to 2649)  
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D., and Cargill, M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 2649)  
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D., and Cargill, M.  
TITLE Direct Subsequencing  
JOURNAL Submitted (18-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source Location/Qualifiers  
1..2649 /organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
gene <1..>2649  
/gene="DPP8"  
/locus\_tag="HCM4283"

Query Match 59.9%; Score 717.4; DB 10; Length 2649;  
Best Local Similarity 81.1%; Pred. No. 3, 4e-187;  
Matches 878; Conservative 0; Mismatches 57; Indels 147; Gaps 1;  
QY 1 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAA 60  
DB 1568 ATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAA 1627  
QY 61 ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGTACTCACATTTCTGTGCATCAGTC 120  
DB 1628 ATCTGGAGAGGTGACAAGGCTGACTGACCGTGGTACTCACATTTCTGTGCATCAGTC 1687  
QY 121 AGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180  
DB 1688 AGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 1747  
QY 181 TTTACAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAAATTTTGGGCCA 240  
DB 1748 TTTACAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAAATTTTGGGCCA 1807  
QY 241 CCATTTGGATTTCAGAGGCTCTCTGACTACTCTCCAGAAATTTCTCTTTG 300  
DB 1808 CCATTTGGATTTCAGAGGCTCTCTGACTACTCTCCAGAAATTTCTCTTTG 1867  
QY 301 AAAGTACTACTGGATTTACATTTGTATGGATCTCTACAAGCTCATGATCTACAGCTG 360  
DB 1868 AAAGTACTACTGGATTTACATTTGTATGGATCTCTACAAGCTCATGATCTACAGCTG 1927  
QY 361 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCA----- 403  
DB 1928 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCAAGTGCAGTGTGGTGAATA 1987  
QY 404 ----- 403

Db 1988 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGGAATACCTTACCTCTCTAGGTATGTGG 2047  
Qy 404 ----- 403  
Db 2048 TTGTAGTGATAGACAACAGGGGATCTGTACCGGAGGCTTAAATTTTGAAGGCGCCTTTA 2107  
Qy 404 -----GGGCTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453  
Db 2108 AATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 2167  
Qy 454 CTCGATGATGATTTTCATGACTTAGATCGGTGGGATCCACGGCTGGTCTTATGGAGGAT 513  
Db 2168 CTCGATGATGATTTTCATGACTTAGATCGGTGGGATCCACGGCTGGTCTTATGGAGGAT 2227  
Qy 514 ACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 573  
Db 2228 ACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 2287  
Qy 574 CCCAGTCTACTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCCACC 633  
Db 2288 CCCAGTCTACTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCCACC 2347  
Qy 634 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCATCGCAAGCAGAAAAAGTTCC 693  
Db 2348 CTGACCAAGATGAACANNN 2407  
Qy 694 CCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTCCATTTTG 753  
Db 2408 NNNNNNACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTCCATTTTG 2467  
Qy 754 CACATACCAATATTTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTACAGA 813  
Db 2468 CACATACCAATATTTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTACAGA 2527  
Qy 814 TCTATCTCGGAGAGACACAGCATAGATTCCTGAATCGGGAGACATATGAATGCTG 873  
Db 2528 TCTATCTCGGAGAGACACAGCATAGATTCCTGAATCTGGAGACATATGAATGCTG 2587  
Qy 874 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCAGTATTCCTGCTCTAAAGTGATAT 933  
Db 2588 ATCTTTTGCACTACCTTCAAGAAACCTTGGATCAGTATTCCTGCTCTAAAGTGATAT 2647  
Qy 934 AA 935  
Db 2648 AA 2649  
RESULT 10  
AK029788  
LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930560C15 product:DIPEPTIDYL PEPTIDASE 8 homolog (Homo sapiens), full insert sequence.  
ACCESSION AK029788  
VERSION AK029788.1 GI:26081520  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
11042159

REFERENCE  
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
11076861

JOURNAL  
PUBMED

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

REFERENCE  
AUTHORS

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
PUBMED

6 (bases 1 to 5517)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/.

FEATURES  
source

1. 5517  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:4930560C15"  
/db\_xref="taxon:10090"  
/clone="4930560C15"  
/sex="male"  
/tissue\_type="testis"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
2221. 4837  
/note="DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens] (SPTRIQ9HBM5, evidence: FASTY, 95.7%ID, 100%length, match=2649)  
putative"

## misc\_feature

2221. 4837  
/note="DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens] (SPTRIQ9HBM5, evidence: FASTY, 95.7%ID, 100%length, match=2649)  
putative"

## ORIGIN

Query Match 58.9%; Score 704.8; DB 4; Length 5517;  
Best Local Similarity 76.1%; Pred. No. 1.3e-183;

Matches 1008; Conservative 0; Mismatches 157; Indels 159; Gaps 5;		
Qy	1	ATTTTGAAGCCCAAGAAGTCCCTTTAGACATCACCTGTAGTGTAGTTCAGTAA 60
Db	3756	ACTTTGAAGCCCAAGAAGTCTCTTTGGAGCATCACCTGTAGTGTAGTTCAGTAA 3815
Qy	61	ATCTGTGAGGTGACAGGCTGACGACGTTGGTGTACTCACTTCTGTGTGACATGTC 120
Db	3816	ACCTGTGAGGTGAGGCTGACGACGTTGGTGTACTCACTTCTGTGTGACATGTC 3875
Qy	121	AGCACTGTGATCTTTTAAAGTAAGTATAGTAAACAGAGAATCCACATGTGTGTC 180
Db	3876	GGCAATTGTGATCTTTTAAAGTAAGTATAGTAAACAGAGAATCCACATGTGTGTC 3935
Qy	181	TTTAAAGCTATCAAGTCTGAGATGACCCCACTTGCAGAAACAAAGGAATTTTGGGCCA 240
Db	3936	TCTAACAATCTCAAGTCTGAGATGACCCCACTTGCAGAAACAAAGGAATTTTGGGCCA 3995
Qy	241	CCATTTTGGATTGACAGGCTCTCTTCTGATATATCTCTCCAGAAATTTTCTTTTG 300
Db	3996	CCATTTTGGATTGACAGGCTCTCTTCTGATATATCTCTCCAGAAATTTTCTTTTG 4055
Qy	301	AAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCTG 360
Db	4056	AAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCTG 4115
Qy	361	GAAGAAATATCTGCTGTGTTTATATATGTTGGTCTCTCA----- 403
Db	4116	GAAGAAATATCTGCTGTGTTTATATATGTTGGTCTCTCA----- 4175
Qy	404	----- 403
Db	4176	ATCGGTTTAAAGAGTCAAGTATTTCCGCGTGAACACCTGGCTCTCTGGTTATGTGG 4235
Qy	404	----- 403
Db	4236	TTGTGTGTATAGACACAGGGATCTGTACCGAGGACTTAATTTGAAGGCGCCTTTA 4295
Qy	404	-----GGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTT 453
Db	4296	AATATAAATGGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCAT 4355
Qy	454	CTGATATAGTTTCTAGTCTAGATCTGTGTGGGATCCAGCGTGTCTCTATGGAGAT 513
Db	4356	CTGATATAGTTTCTAGTCTAGATCTGTGTGGGATCCAGCGTGTCTCTATGGAGAT 4415
Qy	514	ACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTGTCTATTGCTGGGG 573
Db	4416	ACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTGTCTATTGCTGGGG 4475
Qy	574	CCCAGTCTACTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 633
Db	4476	CCCAGTCTACTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 4535
Qy	634	CTGACAGAAATGAACAGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAGAAAGTTCC 693
Db	4536	CTGACAGAAATGAACAGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAGAAAGTTCC 4595
Qy	694	CCTCTGAACCAATCGTTTACTGCTTCTACATGGTTTCTGTGATGAGAATGTCCATTG 753
Db	4596	CCTCTGAACCAATCGTTTACTGCTTCTACATGGTTTCTGTGATGAGAATGTCCATTG 4655
Qy	754	CACATACAGTATATTACTGATTTTTTTAGTGGGCTGGAAAGCCATATGATTACAGA 813
Db	4656	CACATACAGTATATTACTGATTTTTTTAGTGGGCTGGAAAGCCATATGATTACAGA 4715
Qy	814	TCTATCTCAGGAGACACACATAGATTTCTGTATCGGAGACATTTAGTACTGC 873
Db	4716	TCTATCTCAGGAGACACACATAGATTTCTGTATCGGAGACATTTAGTACTGC 4775
Qy	874	ATCTTTTGGACTCTTCAAGAAACCTTGGATCAGCTATTGCTCTTAAGATCATAT 933
Db	4776	ATCTTTTGGACTCTTCAAGAAACCTTGGATCAGCTATTGCTCTTAAGATCATAT 4835

Qy	934	AAATTTGACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTATTAACCAAAATGAGAGGTT 993
Db	4836	AACTCTGACCATGTAGAACT---TGGTGTACACTG---TCTAACCAAAACAAAGAGGTT 4888
Qy	994	TAATCAACAGAAACACAGAAATGTGATCATCATTTTGTATACCTGCTGCAATGATCTAC 1053
Db	4889	TAGCAATAGAGAGCCAGAGTTGATGAACCTTACTTTTGTACCTGCTGCAATGATCTAT 4948
Qy	1054	TCTGAAAATATAATTTGGTGGCCATGACGGGTCTACGGTTTGTGG-TAGTAAATCTAATAC 1112
Db	4949	TTCTGAAGTAAATTTGGTGGCCATGACAAATCTGCAGTTTCATGTATGTACTCTTAATAC 5008
Qy	1113	CTTAACCCCATGCTCAAAATGATACATATTTCTGTAGAGACCCAGCAATACCAT 1172
Db	5009	CTTAAC---CATGCACATAAACTGAGTGACATGTTCCCAAGGATGAACAGTACCTG 5064
Qy	1173	AAGA 1176
Db	5065	AGGA 5068
RESULT 11		
AK016546		
LOCUS	3143 bp	mRNA linear HTC 03-APR-2004
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932434F09 product:DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens], full insert sequence.	
ACCESSION	AK016546	1 GI:12855334
VERSION	AK016546	1
KEYWORDS	HTC; CAP trapper.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)	
AUTHORS	Carninci, P. and Hayashizaki, Y.	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	
PUBMED	11042159	
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)	
JOURNAL	11076861	
PUBMED	11076861	
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)	
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase II Team.	
TITLE	Analysis of the mouse transcriptome based on functional annotation of 6,770 full-length cDNAs	
JOURNAL	Nature 420, 563-573 (2002)	
PUBMED	6 (bases 1 to 3143)	
REFERENCE	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,	



```
Db      2944  ACCTGCTCCACTACCTTCAGGAGAACCTTGGATCGGTATCGTCTCTGAAAGTGATAT 3003
Qy      934  AATTTTACCTGTTAGAACTCTCTGGTATACACTGGCTATTTAACCAAAATGAGGAGTT 993
Db      3004  AACTCTGACGATGTAGAACT---TGGTGTACACTG---TCTAACCAAAACAAGAGGGT 3056
Qy      994  TAATCAACAAGAAACACAGAAATGATCATCAATTTTGTATACCTGGCCATGTACATCTAC 1053
Db      3057  TAGCCATGAGAGGACGAGTTGATGAACCTACTTTTGTACCTGCCACATAGAACTAT 3116
Qy      1054  TCTGAAAATTAATGTGTGCCATGC 1079
Db      3117  TTCTGAAGTAATTTTGGTGCCATGC 3142

RESULT 12
LOCUS   AL043338                735 bp      mRNA      linear      EST 04-SEP-2003
DEFINITION DRPZp43400723_v1_434 (synonym: htes3) Homo sapiens cDNA clone
AL043338
VERSION  AL043338.1 GI:5422728
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 735)
          Blum,H., Baueraachs,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
          EST (Blum, et al.)
          Unpublished (1999)
          Contact: MIPS
          MIPS
          Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
          This is the 5' sequence of the clone insert.
          Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
          Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
          sequenced by LMU (Ludwig Maximilians University,
          Munich/Germany) within the cDNA sequencing consortium of the German
          Genome Project.
          No si sequence available.
          This clone (DRPZp43400723) is available at the RZPD in Berlin.
          Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
          Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
          Location/Qualifiers
            1..735
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="DRPZp43400723"
              /tissue_type="testis"
              /dev_stage="adult"
              /lab_host="DH10B"
              /clone_lib="434 (synonym: htes3)"
              /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

FEATURES             source
  Query Match          56.6%; Score 677; DB 1; Length 735;
  Best Local Similarity 99.9%; Pred. No. 3.8e-176;
  Matches 677; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      404  GGGTCAATAGAAATGACGATCAGGTGGAGGACTCCATATCTAGCTTCGATATCA 463
Db      58  GGGTCAATAGAAATGACGATCAGGTGGAGGACTCCATATCTAGCTTCGATATCA 117
Qy      464  TTTCATTGACTTAGATCGTGTGGGATCCACGGCTGCTCTATGGAGGATACCTTCCTCT 523
Db      118  TTTCATTGACTTAGATCGTGTGGGATCCACGGCTGCTCTATGGAGGATACCTTCCTCT 177
Qy      524  GATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 583
```

```
Db      178  GATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 237
Qy      584  TCTGTGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTACCTGACCAGAA 643
Db      238  TCTGTGATCTTCTATGATACAGGATACACGGAAGCTTATATGGGTACCTGACCAGAA 297
Qy      644  TGAACAGGGCTATTAATCTAGGATCTGTGGCCATGACAGAGAAAAGTTCCCTCTGAAAC 703
Db      298  TGAACAGGGCTATTAATCTAGGATCTGTGGCCATGACAGAGAAAAGTTCCCTCTGAAAC 357
Qy      704  AAATCGTTTACTGCTCTTACATGGTTTCTGTGATGAGATGTCCATTTTGGACATPACCA 763
Db      358  AAATCGTTTACTGCTCTTACATGGTTTCTGTGATGAGATGTCCATTTTGGACATPACCA 417
Qy      764  TATAATTAATGAGTTTATAGTGGGCTGGAAGGCATATGATTTACAGATCTATCTCTCA 823
Db      418  TATAATTAATGAGTTTATAGTGGGCTGGAAGGCATATGATTTACAGATCTATCTCTCA 477
Qy      824  GGAGAGACACAGCATAGAGTTTCCCTGAATCGGGAGAACATTAATGAACTGTCATCTTTGCA 883
Db      478  GGAGAGACACAGCATAGAGTTTCCCTGAATCGGGAGAACATTAATGAACTGTCATCTTTGCA 537
Qy      884  CTACCTTTCAACAAAACCTTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACC 943
Db      538  CTACCTTTCAACAAAACCTTTGGATCAGTATTCCTCTCTAAAGTGATATAATTTTGACC 597
Qy      944  TGTGTAGAACTCTCTGTATACATCGGCTATTTTACCAAAATGAGGAGGTTTAATCAACAG 1003
Db      598  TGTGTAGAACTCTCTGTATACATCGGCTATTTTACCAAAATGAGGAGGTTTAATCAACAG 657
Qy      1004  AAAACACAGAAATGATCAATTCATGATTTGATACCTGCGCATGAATGAACTGTCATCTCTGAA 1063
Db      658  AAAACACAGAAATGATCAATTCATGATTTGATACCTGCGCATGAATGAACTGTCATCTCTGAA 717
Qy      1064  AAATGTGGTGCCATGCAAG 1081
Db      718  AAATGTGGTGCCATGCAAG 735

RESULT 13
LOCUS   BG479035
DEFINITION 602526111F1 NTH_MGC_21 Homo sapiens cDNA clone IMAGE:4649834 5',
BG479035
VERSION  BG479035
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 910)
          NIH-MGC http://mgs.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          TITLE
          JOURNAL
          COMMENT
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.lnl.gov
          Plate: LNCMI430 row: e column: 03
          High quality sequence stop: 837.
          Location/Qualifiers
            1..910
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:4649834"

FEATURES             source
  Query Match          56.6%; Score 677; DB 1; Length 735;
  Best Local Similarity 99.9%; Pred. No. 3.8e-176;
  Matches 677; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      404  GGGTCAATAGAAATGACGATCAGGTGGAGGACTCCATATCTAGCTTCGATATCA 463
Db      58  GGGTCAATAGAAATGACGATCAGGTGGAGGACTCCATATCTAGCTTCGATATCA 117
Qy      464  TTTCATTGACTTAGATCGTGTGGGATCCACGGCTGCTCTATGGAGGATACCTTCCTCT 523
Db      118  TTTCATTGACTTAGATCGTGTGGGATCCACGGCTGCTCTATGGAGGATACCTTCCTCT 177
Qy      524  GATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 583
```

/tissue\_type="choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NH\_MGC\_21"  
/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 56.3%; Score 674.2; DB 2; Length 910;  
Best Local Similarity 94.7%; Pred. No. 2.4e-175;  
Matches 730; Conservative 0; Mismatches 38; Indels 3; Gaps 3;

QY 404 GGTCAAAATAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 463  
DB 139 GGGTCAAAATAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 198  
QY 464 TTTTCATTGACTTGTAGTTCGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCT 523  
DB 199 TTTTCATTGACTTGTAGTTCGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCT 258  
QY 524 GATGGCATTAAATGCAGAGGTGATATCTTCAGGGTGTCTATGCTGGGGCCCCAGTCA 583  
DB 259 GATGGCATTAAATGCAGAGGTGATATCTTCAGGGTGTCTATGCTGGGGCCCCAGTCA 318  
QY 584 TCTGTGGATCTTCTATGATACAGGATACACGGAGCTTATATGGGTCACTCCCTGACCAGAA 643  
DB 319 TCTGTGGATCTTCTATGATACAGGATACACGGAGCTTATATGGGTCACTCCCTGACCAGAA 378  
QY 644 TGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGACC 703  
DB 379 TGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGACC 438  
QY 704 AATCTGTTACTGCTTCTACATGTTTCTCTGATGAGAGTCTCCATTTTGCATATCTCA 823  
DB 439 AATCTGTTACTGCTTCTACATGTTTCTCTGATGAGAGTCTCCATTTTGCATATCTCA 498  
QY 764 TATATTACTGAGTCTTCTAGTGGCTGGAAGCCATATGTTTACAGATCTATCTCA 823  
DB 499 TATATTACTGAGTCTTCTAGTGGCTGGAAGCCATATGTTTACAGATCTATCTCA 558  
QY 824 GGAGACACAGCATAGAGTCTCTGAATCGGAGAACATTTATGAATGATCTTTTGA 883  
DB 559 GGAGACACAGCATAGAGTCTCTGAATCGGAGAACATTTATGAATGATCTTTTGA 618  
QY 884 CTACCTTCAAGAAAACCTTGGATCAGTATGCTGCTCTAAAGTGTATATAATTTGACC 943  
DB 619 CTACCTTCAAGAAAACCTTGGATCAGTATGCTGCTCTAAAGTGTATATAATTTGACC 678  
QY 944 TGTGTAGACTCTCTGGTATACATCTGGTCTTATTAACAAATGAGGAGTTTATCAACAG 1003  
DB 679 TGTGTAGACTCTCTGGTATACATCTGGTCTTATTAACAAATGAGGAGTTTATCAACAG 738  
QY 1004 AAAACACAGAAATGATCATCATTTTGTATCTCTCCATGTAAATCTACTCTCTGAAAT 1063  
DB 739 GAAACACAGAAATGATCATCATTTTGTATCTCTCCATGTAAATCTACTCTCTGAAAT 797  
QY 1064 AA-ATGTGTGCCATCGAGGGTCTACGGTTTGTGGTAGTAAATCTTAATACCTTAACCCCA 1122  
DB 798 AATGTGGTGGCCATCGAGGGTCTACGGGAGTGAAGTAATCAATATCTTTAAACCCGA 857  
QY 1123 CATGCTCAAAATCAATGATACATATTTCTGTAGAGACCCAGCAATACCATTA 1173  
DB 858 AATGGTTCAATCAATTTGGT-CATTTTCTGAGAAAACAGCAGTAAACCATTA 907

RESULT 14  
BU622228/c  
LOCUS BU622228 708 bp mRNA linear EST 23-SEP-2002

DEFINITION UI-H-FHI-bfu-d-21-0-UI.s1 NCI CGAP FHI Homo sapiens cDNA clone  
UI-H-FHI-bfu-d-21-0-UI 3', mRNA sequence.  
ACCESSION BU622228 GI:23288443  
VERSION BU622228.1  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 708)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITILE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: James Martin  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, bento-soares@uiowa.edu  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
source  
1..708  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-FHI-bfu-d-21-0-UI"  
/tissue\_type="Cell Line"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP FHI"  
/note="Organ: Chondrosarcoma; Vector: p773-Pac  
(Pharmacia) with a modified polylinker; Site 1: EcoR I;  
Site 2: Not I; NCI CGAP FHI is a normalized cDNA library  
obtained from a cell line derived from grade I  
chondrosarcoma tissue. The library was constructed and  
normalized according to Bonaldo, Lennon and Soares, Genome  
Research, 6:791-806, 1996. First strand cDNA synthesis was  
primed with an oligo-dT primer containing a Not I site.  
Double stranded cDNA was ligated to an EcoR I adaptor,  
digested with Not I, and cloned directionally into  
p773-Pac vector. The oligonucleotide used to prime the  
synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
clone distribution information. The sequence tag for this library is  
(dt)18 tail. The sequence tag for this library is  
AGAATCCGGC. The cell line was provided by Dr. James Martin  
from the University of Iowa  
TAG\_TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1  
Chondrosarcoma  
TAG\_LIB=UI-H-FHI  
TAG\_SEQ=AGAATCCGGC"

ORIGIN

Query Match 56.3%; Score 673.8; DB 5; Length 708;  
Best Local Similarity 99.7%; Pred. No: 2.9e-175;  
Matches 675; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 404 GGGTCAAAATAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 463  
DB 693 GGGTCAAAATAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGA 634  
QY 464 TTTTCATTGACTTGTAGTTCGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCT 523  
DB 633 TTTTCATTGACTTGTAGTTCGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCT 574  
QY 524 GATGGCATTAAATGCAGAGGTGATATCTTCAGGGTGTCTATGCTGGGGCCCCAGTCA 583  
DB 573 GATGGCATTAAATGCAGAGGTGATATCTTCAGGGTGTCTATGCTGGGGCCCCAGTCA 514



QY 584 TCTGTGGATCTTCTATGATACAGGATACACGGAAGTTATATGGGTCACTGACCGAGAA 643  
Db 513 TCTGTGGATCTTCTATGATACAGGATACACGGAAGTTATATGGGTCACTGACCGAGAA 454  
QY 644 TGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACC 703  
Db 453 TGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGHACC 394  
QY 704 AATATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGGACATACAG 763  
Db 393 AATATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGGACATACAG 334  
QY 764 TATATTACTAGTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA 823  
Db 333 TATATTACTAGTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCA 274  
QY 824 GGAGAGACAGCATTAAGAGTTCTTGAATCGGGAGAACATTTATGAATGCTATCTTTTGA 883  
Db 273 GGAGAGACAGCATTAAGAGTTCTTGAATCGGGAGAACATTTATGAATGCTATCTTTTGA 214  
QY 884 CTACCTTCAAGAAACCTTTGGATCAGCTATTGCTCTTAAAGTGATATAATTTTGACC 943  
Db 213 CTACCTTCAAGAAACCTTTGGATCAGCTATTGCTCTTAAAGTGATATAATTTTGACC 154  
QY 944 TGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAG 1003  
Db 153 TGTGTAGAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAG 94  
QY 1004 AAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAACATCTCTCTGAAAT 1063  
Db 93 AAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAACATCTCTCTGAAAT 34  
QY 1064 AAATGTTGGCCATGCA 1080  
Db 33 AAATGTTGGCCATGNA 17

RESULT 15  
LOCUS AW303607/c 673 bp mRNA linear EST 18-JAN-2000  
DEFINITION xz1b12.x1 Soares\_NFL\_T\_GBC\_S1 Homo sapiens cDNA clone  
IMAGE:2813759 3' similar to TR:075273 075273 R26984\_1 ;, mRNA  
sequence.  
ACCESSION AW303607 GI:6713296  
VERSION AW303607.1  
KEYWORDS EST.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
1 (bases 1 to 673)  
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -40UP from Gibco  
High quality sequence stop: 465.  
FEATURES  
Location/Qualifiers  
1..673  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2813759"  
/lab\_host="DH10B"  
/clone\_libs="Soares\_NFL\_T\_GBC\_S1"  
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;  
Equal amounts of plasmid DNA from three normalized

libraries (fetal lung NbHL19W, testis NHT, and B-cell  
NCI-CGAP GCBI) were mixed, and ss circles were made in  
vivo. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of  
I.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo. "

## ORIGIN

Query Match 55.9%; Score 668.8; DB 1; Length 673;  
Best Local Similarity 99.6%; Pred. No. 7e-174;  
Matches 670; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 510 GGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCT 569  
Db 673 GGATACCTCTCCNTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCT 614  
QY 570 GGGGCCCGACGTCACTCTGTGGATCTTCTATGATACAGGATACAGGACGTTATATGGGT 629  
Db 613 GGGGCCCGACGTCACTCTGTGGATCTTCTATGATACAGGATACAGGACGTTATATGGGT 554  
QY 630 CACCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCAGTGGAGGAGAAAG 689  
Db 553 CACCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCAGTGGAGGAGAAAG 494  
QY 690 TTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGATGTCAT 749  
Db 493 TTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGATGTCAT 434  
QY 750 TTTGCAATACCAAGTATATTTACTGAGTTTCTGAGGGCTGGAAGCCATATGATTTA 809  
Db 433 TTTGCAATACCAAGTATATTTACTGAGTTTCTGAGGGCTGGAAGCCATATGATTTA 374  
QY 810 CAGATCTATCTCAGAGAGACACAGCAATAGAGTTCTGAAATCGGAGAACATTATGAA 869  
Db 373 CAGATCTATCTCAGAGAGACACAGCAATAGAGTTCTGAAATCGGAGAACATTATGAA 314  
QY 870 CTGATCTTTTGCATCTACCTTCAAGAAAACCTTTGGATCACGTTATGCTGCTCTAAAAGTG 929  
Db 313 CTGATCTTTTGCATCTACCTTCAAGAAAACCTTTGGATCACGTTATGCTGCTCTAAAAGTG 254  
QY 930 ATATAATTTTGACCTGTGTAGAACTCTCTGATATACACTGGCTATTTAAACCAATGAGGA 989  
Db 253 ATATAATTTTGACCTGTGTAGAACTCTCTGATATACACTGGCTATTTAAACCAATGAGGA 194  
QY 990 GGTTTAATCAACAGAAAACACAGAAATTCATCATCATTTTGTATACCTGCAATGTAACAT 1049  
Db 193 GGTTTAATCAACAGAAAACACAGAAATTCATCATCATTTTGTATACCTGCAATGTAACAT 134  
QY 1050 CTACTCTGAAAATAAATGTGTGCGCATGCAAGGGTCTACGGTTTGTAGTAATCTAA 1109  
Db 133 CTACTCTGAAAATAAATGTGTGCGCATGCAAGGGTCTACGGTTTGTAGTAATCTAA 74  
QY 1110 TACCTTAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGACCCAGCAATAC 1169  
Db 73 TACCTTAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGACCCAGCAATAC 14  
QY 1170 CATAGAAATTAAT 1182  
Db 13 CATAGAAATTAAT 1

Search completed: May 4, 2006, 07:56:18  
Job time : 4493.98 secs



Result No.	Score	Query Match	Length	DB	ID	SUMMARIES	
						%	Description
1	1197	100.0	1197	3	US-10-070-464-4	Sequence 4, Appl	
2	1179	98.5	4676	3	US-09-976-674-20	Sequence 20, Appl	
3	1040	86.9	2797	3	US-09-976-594-1103	Sequence 1103, Appl	
4	1036	86.5	3120	3	US-10-070-464-2	Sequence 2, Appl	
5	1033.2	86.3	4309	3	US-09-976-674-14	Sequence 14, Appl	
6	1033.2	86.3	4829	3	US-09-976-674-12	Sequence 12, Appl	
7	1013.6	84.7	4685	3	US-09-976-674-22	Sequence 22, Appl	
8	881	73.6	1669	3	US-10-070-464-6	Sequence 6, Appl	
9	873	72.9	4523	3	US-09-976-674-8	Sequence 8, Appl	
10	776.4	64.9	2671	3	US-09-976-674-2	Sequence 2, Appl	
11	633.4	52.9	8271	3	US-09-280-116-171	Sequence 171, Appl	
12	459	38.3	1083	3	US-10-070-464-8	Sequence 8, Appl	
13	326.8	27.3	612	3	US-09-962-184-31	Sequence 31, Appl	
14	297	24.8	308	3	US-09-621-976-1787	Sequence 1787, Appl	
15	212.8	17.8	2617	3	US-09-976-674-4	Sequence 4, Appl	
16	212.8	17.8	4180	3	US-09-976-674-36	Sequence 36, Appl	
17	212.8	17.8	4219	3	US-09-976-674-28	Sequence 28, Appl	
18	212.8	17.8	4263	3	US-09-976-674-34	Sequence 34, Appl	
19	212.8	17.8	4302	3	US-09-976-674-24	Sequence 24, Appl	
20	182.6	15.3	4076	3	US-09-976-674-32	Sequence 32, Appl	
21	182.6	15.3	4159	3	US-09-976-674-30	Sequence 30, Appl	
22	181.4	15.2	4037	3	US-09-976-674-40	Sequence 40, Appl	
23	181.4	15.2	4120	3	US-09-976-674-38	Sequence 38, Appl	
24	86	7.2	889	3	US-09-513-999C-16433	Sequence 16433, Appl	

Db 241 CCATTTTGGATTACAGAGTCTCTTCTGACTATACCTCCACGAAATTTCTCTTTTG 300  
QY 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCTG 360  
Db 301 AAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCTG 360  
QY 361 GAAAGAAATATCTTACTGCTGCTTTATATATGTTGGTCTCTCAGGGTCAAAATGAATG 420  
Db 361 GAAAGAAATATCTTACTGCTGCTTTATATATGTTGGTCTCTCAGGGTCAAAATGAATG 420  
QY 421 ACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATC 480  
Db 421 ACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATC 480  
QY 481 GTGFGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGCAATTAATGAGA 540  
Db 481 GTGFGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGCAATTAATGAGA 540  
QY 541 GGTGAGATATCTTACGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 600  
Db 541 GGTGAGATATCTTACGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 600  
QY 601 ATACAGGATACAGGAACTTTATATGGGTCACTCTGAGCAGAGAAAGAGGCTATTACT 660  
Db 601 ATACAGGATACAGGAACTTTATATGGGTCACTCTGAGCAGAGAAAGAGGCTATTACT 660  
QY 661 TAGGATCTGTGGCATGCAAGCAAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 720  
Db 661 TAGGATCTGTGGCATGCAAGCAAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 720  
QY 721 TACATGTTTCTCTGGATGAGATGTCATTTTGGACATACCATGATATATTACTAGTTTTT 780  
Db 721 TACATGTTTCTCTGGATGAGATGTCATTTTGGACATACCATGATATATTACTAGTTTTT 780  
QY 781 TAGTGAGGCTGGAGGCAATGATTTACAGATCTATCTCCAGAGAGACACAGCATAA 840  
Db 781 TAGTGAGGCTGGAGGCAATGATTTACAGATCTATCTCCAGAGAGACACAGCATAA 840  
QY 841 GAGTTCCTGAACTGGAGAACATATGAACTGCTCTTTTGGACATCTCTTCAAGAAACC 900  
Db 841 GAGTTCCTGAACTGGAGAACATATGAACTGCTCTTTTGGACATCTCTTCAAGAAACC 900  
QY 901 TTGGATCAGCTATGCTGCTCTAAAGTGATATATTTTGAACCTGTGTAGAACCTCTG 960  
Db 901 TTGGATCAGCTATGCTGCTCTAAAGTGATATATTTTGAACCTGTGTAGAACCTCTG 960  
QY 961 TATACATGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATC 1020  
Db 961 TATACATGCTATTTAAACAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATC 1020  
QY 1021 ATACATTTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
Db 1021 ATACATTTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
QY 1081 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAATG 1140  
Db 1081 GGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAATG 1140  
QY 1141 ATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAAATAAAAA 1197  
Db 1141 ATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAAATAAAAA 1197

RESULT 2  
US-09-976-674-20  
; Sequence 20, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PENDING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PENDING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 4676  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-20  
  
Query Match 98.5%; Score 1179; DB 3; Length 4676;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1193; Conservative 0; Mismatches 0; Indels 4; Gaps 1;  
  
QY 1 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACAGTTACGTAA 60  
Db 1781 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACAGTTACGTAA 1840  
QY 61 ATCTGTGAGAGGTGACAAAGGCTGACTGACGCTGGCTTACTCACAATTTCTTGTGATCAGTC 120  
Db 1841 ATCTGTGAGAGGTGACAAAGGCTGACTGACGCTGGCTTACTCACAATTTCTTGTGATCAGTC 1900  
QY 121 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCC 180  
Db 1901 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCC 1960  
QY 181 TTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGAAAAACAAGGAATTTTGGGCA 240  
Db 1961 TTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGGAAAAACAAGGAATTTTGGGCA 2020  
QY 241 CCATTTTGAATTCAGAGGCTCTCTCTGATGATATCTCTCCAGAAATTTCTCTTTG 300  
Db 2021 CCATTTTGAATTCAGAGGCTCTCTCTGATGATATCTCTCCAGAAATTTCTCTTTG 2080  
QY 301 AAGTACTGCTGATTTACATTTGATGGATGCTCTCAAGGCTCATGATCTACAGCTG 360  
Db 2081 AAGTACTGCTGATTTACATTTGATGGATGCTCTCAAGGCTCATGATCTACAGCTG 2140  
QY 361 GAAAGAAATCTCTACTGCTGCTTCAATATATGTTGGTCTCTCAGGGTCAAAATAGAAATG 420  
Db 2141 GAAAGAAATCTCTACTGCTGCTTCAATATATGTTGGTCTCTCAGGGTCAAAATAGAAATG 2196  
QY 421 ACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCACTTAGATC 480  
Db 2197 ACGATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCACTTAGATC 2256  
QY 481 GTGFGGCATCCACGGCTGGTCTATGAGGAGATACCTCTCCCTGATGCAATTAATGAGA 540  
Db 2257 GTGFGGCATCCACGGCTGGTCTATGAGGAGATACCTCTCCCTGATGCAATTAATGAGA 2316  
QY 541 GGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 600  
Db 2317 GGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATG 2376  
QY 601 ATACAGGATACAGGAACTTTATATGGGTCACTCTGAGCAGAGAAAGAGGCTATTACT 660  
Db 2377 ATACAGGATACAGGAACTTTATATGGGTCACTCTGAGCAGAGAAAGAGGCTATTACT 2436  
QY 661 TAGGATCTGTGGCATGCAAGCAAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 720  
Db 2437 TAGGATCTGTGGCATGCAAGCAAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCT 2496  
QY 721 TACATGTTTCTCTGGATGAGATGTCATTTTGGACATACCATGATATATTACTAGTTTTT 780  
Db 2497 TACATGTTTCTCTGGATGAGATGTCATTTTGGACATACCATGATATATTACTAGTTTTT 2556  
QY 781 TAGTGAGGCTGGAGGCAATGATTTTACAGATCTCTCTCAGAGAGACACAGCATAA 840  
Db 2557 TAGTGAGGCTGGAGGCAATGATTTTACAGATCTCTCTCAGAGAGACACAGCATAA 2616

```
QY 841 GAGTTCCTGAATCGGAGAACATTAATGAACCTGATCTTTTGACCTACCTTGAAGAAACC 900
Db 2617 GAGTTCCTGAATCGGAGAACATTAATGAACCTGATCTTTTGACCTACCTTCAAGAAACC 2676
QY 901 TTGGATCAGTATTGCTGCTCTAAAGTGATATAATTTGACCTGTGTAGAACTCTCTGG 960
Db 2677 TTGGATCAGTATTGCTGCTCTAAAGTGATATAATTTGACCTGTGTAGAACTCTCTGG 2736
QY 961 TATACACTGGCTATTATTAACCAATGAGGAGGTTTAATCAACAGAGAAACACAGAATTGATC 1020
Db 2737 TATACACTGGCTATTATTAACCAATGAGGAGGTTTAATCAACAGAGAAACACAGAATTGATC 2796
QY 1021 ATCACATTTTGAATACCTGCCATGTAAACATCTACTCTCTGAAATAAATGTTGGCCATGCA 1080
Db 2797 ATCACATTTTGAATACCTGCCATGTAAACATCTACTCTCTGAAATAAATGTTGGCCATGCA 2856
QY 1081 GGGGCTACGGTTTGTGTAGTAACTTAATACCTTAACCCACATGCTCAAAATCAAATG 1140
Db 2857 GGGGCTACGGTTTGTGTAGTAACTTAATACCTTAACCCACATGCTCAAAATCAAATG 2916
QY 1141 ATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTACTTAAAAAAA 1197
Db 2917 ATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTACTTAAAAAAA 2973

RESULT 3
US-09-976-594-1103
; Sequence 1103, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 1103
; LENGTH: 2797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1
US-09-976-594-1103

Query Match 86.9%; Score 1040; DB 3; Length 2797;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTTGAAGGACCAAGACTCCCTTTAGAGATCAACCTGTACGTAGTACGTACGTAA 60
Db 1452 ATTTTGAAGGACCAAGACTCCCTTTAGAGATCAACCTGTACGTAGTACGTACGTAA 1511
QY 61 ATCCGTGGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGTC 120
Db 1512 ATCCGTGGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGTC 1571
QY 121 AACTGTGACCTCTTTTATAGTAAAGTATAGTAACAGAGAAATCCACATCTGTGTCCC 180
Db 1572 AACTGTGACCTCTTTTATAGTAAAGTATAGTAACAGAGAAATCCACATCTGTGTCCC 1631
QY 181 TTTCACAGTATCAAGTCTGAGATGACCCAACTGTGCAAAACAAAGGAATTTTGGGCA 240
Db 1632 TTTCACAGTATCAAGTCTGAGATGACCCAACTGTGCAAAACAAAGGAATTTTGGGCA 1691
QY 241 CCATTTTGGATTACGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTG 300
Db 1692 CCATTTTGGATTACGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTG 1751
```

```
QY 301 AAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 360
Db 1752 AAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTG 1811
QY 361 GAAAGAAATATCCTACTGCTGCTTATATATATGCTGCTCTCA----- 403
Db 1812 GAAAGAAATATCCTACTGCTGCTTATATATATGCTGCTCTCA----- 1871
QY 404 ----- 403
Db 1872 ATCCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGCCTCTCTAGGTTATGTGG 1931
QY 404 ----- 403
Db 1932 TTGTAGTATAGACAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTGAAGCGCCTTTA 1991
QY 404 -----GGGTCAAAATAGAAATTAACGATCAGGTGGAAGGACTCCCAATATCTAGCTT 453
Db 1992 AATATAAATGGGTCAAAATAGAAATTAACGATCAGGTGGAAGGACTCCCAATATCTAGCTT 2051
QY 454 CTCGATATGATTTCAATTGACTTATGATCGTGTGGGCAATCCACGGCTGGTCTTATGAGGAT 513
Db 2052 CTCGATATGATTTCAATTGACTTATGATCGTGTGGGCAATCCACGGCTGGTCTTATGAGGAT 2111
QY 514 ACCTCTCCCTGATGGCAATTAATGACAGGTCAGATATCTTCAGGGTTCCTATGCTGGG 573
Db 2112 ACCTCTCCCTGATGGCAATTAATGACAGGTCAGATATCTTCAGGGTTCCTATGCTGGG 2171
QY 574 CCCAGTCACCTCTGCTGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACC 633
Db 2172 CCCAGTCACCTCTGCTGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACC 2231
QY 634 CTGACCAATGAACAGGGCTTATTAATGAGTCTCTGGCCATCGAAGCAGAAAAAGTTCC 693
Db 2232 CTGACCAATGAACAGGGCTTATTAATGAGTCTCTGGCCATCGAAGCAGAAAAAGTTCC 2291
QY 694 CCTCTGAACCAATCGTTTATCTGCTCTTACATGGTTTCTGATGAGAAATGTCATTTTG 753
Db 2292 CCTCTGAACCAATCGTTTATCTGCTCTTACATGGTTTCTGATGAGAAATGTCATTTTG 2351
QY 754 CACATACCAGTATATTAATCTGATTTTTTATGAGGGCTCGAAGCCATATGATTTACAGA 813
Db 2352 CACATACCAGTATATTAATCTGATTTTTTATGAGGGCTCGAAGCCATATGATTTACAGA 2411
QY 814 TCTATCTCTCAGGAGAGACACAGCATAAGGTTCTCTGAATCGGAGAACATTTATGAATGC 873
Db 2412 TCTATCTCTCAGGAGAGACACAGCATAAGGTTCTCTGAATCGGAGAACATTTATGAATGC 2471
QY 874 ATCTTTTGCATCTACTCTTCAAGAAAACTTGGATCAGTATTTGCTGCTCTAAAAGTGATAT 933
Db 2472 ATCTTTTGCATCTACTCTTCAAGAAAACTTGGATCAGTATTTGCTGCTCTAAAAGTGATAT 2531
QY 934 AATTTTGAACCTCTGATGACCTCTCTGATATACACTGGCTTATTAACCAATGAGGAGTT 993
Db 2532 AATTTTGAACCTCTGATGACCTCTCTGATATACACTGGCTTATTAACCAATGAGGAGTT 2591
QY 994 TAAATCAACAGAAAAACACAGAAATGATCAACATTTTGTATGATCTGCAATGTAACATCTAC 1053
Db 2592 TAAATCAACAGAAAAACACAGAAATGATCAACATTTTGTATGATCTGCAATGTAACATCTAC 2651
QY 1054 TCCTGAAATAAATGTTGGTCCATGACGGGCTACAGGTTTGTGGTAGTAATCTAATACC 1113
Db 2652 TCCTGAAATAAATGTTGGTCCATGACGGGCTACAGGTTTGTGGTAGTAATCTAATACC 2711
QY 1114 TTAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGCCCGCAATACCATA 1173
Db 2712 TTAACCCACATGCTCAAAATCAAAATGATATATTTCTGAGAGCCCGCAATACCATA 2771
QY 1174 AGAATTTACTAAAAA 1197
Db 2772 AGAATTTACTAAAAA 2795
```

RESULT 4	
US-10-070-464-2	
; Sequence 2, Application US/10070464	
; Patent No. 6881564	
; GENERAL INFORMATION:	
; APPLICANT: ABBOTT, Catherine Anne	
; APPLICANT: GORRELL, Mark Douglas	
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES	
; FILE REFERENCE: GH-007	
; CURRENT APPLICATION NUMBER: US/10/070,464	
; PRIOR FILING DATE: 2002-03-07	
; PRIOR APPLICATION NUMBER: PCT/AU00/01085	
; PRIOR FILING DATE: 2000-09-11	
; PRIOR APPLICATION NUMBER: AU PQ5709	
; PRIOR FILING DATE: 2000-02-18	
; PRIOR APPLICATION NUMBER: AU PQ2762	
; PRIOR FILING DATE: 1999-09-10	
; NUMBER OF SEQ ID NOS: 8	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 2	
; LENGTH: 3120	
; TYPE: DNA	
; ORGANISM: Homo Sapiens	
US-10-070-464-2	
Query Match	
Best Local Similarity 86.5%; Score 1036; DB 3; Length 3120;	
Matches 1193; Conservative 0; Mismatches 0; Indels 147; Gaps 1;	
Qy	1
Db	1781
Qy	61
Db	1841
Qy	121
Db	1901
Qy	181
Db	1961
Qy	241
Db	2021
Qy	301
Db	2081
Qy	361
Db	2141
Qy	404
Db	2201
Qy	404
Db	2261
Qy	404
Db	2321
Qy	454
Db	2381

Qy	514
Db	2441
Qy	574
Db	2501
Qy	634
Db	2561
Qy	694
Db	2621
Qy	754
Db	2681
Qy	814
Db	2741
Qy	874
Db	2801
Qy	934
Db	2861
Qy	994
Db	2921
Qy	1054
Db	2981
Qy	1114
Db	3041
Qy	1174
Db	3101

RESULT 5	
US-09-976-674-14	
; Sequence 14, Application US/09976674	
; Patent No. 6844180	
; GENERAL INFORMATION:	
; APPLICANT: Qi, Steve	
; APPLICANT: Akinsanya, Karen	
; APPLICANT: Riviere, Pierre	
; APPLICANT: Junien, Jean-Louis	
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV	
; FILE REFERENCE: 70669	
; CURRENT APPLICATION NUMBER: US/09/976,674	
; PRIOR FILING DATE: 2001-10-12	
; PRIOR APPLICATION NUMBER: US 60/240,117	
; PRIOR FILING DATE: 2000-10-12	
; NUMBER OF SEQ ID NOS: 61	
; SOFTWARE: PatentIn version 3.1	
; SEQ ID NO 14	
; LENGTH: 4309	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-09-976-674-14	
Query Match	
86.3%; Score 1033.2; DB 3; Length 4309;	

Db	2281	GCATCTTTTGGCACTACCTTTCAAGAAACCTTGGATCAGCTATTTCCTCTCTAAAAGTGAT	2344
Qy	932	ATAATTTTGAAGGCTGTGTAGAACTCTCTGTGTATATACCTGGCTATTTTAAACCAATGAGGAGG	991
Db	2341	ATAATTTTGAAGGCTGTGTAGAACTCTCTGTGTATATACCTGGCTATTTTAAACCAATGAGGAGG	2400
Qy	992	TTTAAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAAACATCT	1051
Db	2401	TTTAAATCAACAGAAACACAGAAATGATCATCATTTTGTATACCTGCCATGTAAACATCT	2460
Qy	1052	ACTCTCTGAAATAAATGTGTGTCATGCAGGGGTCTACGGTTTCTGTGTAGTAACTTAATA	1111
Db	2461	ACTCTCTGAAATAAATGTGTGTCATGCAGGGGTCTACGGTTTCTGTGTAGTAACTTAATA	2520
Qy	1112	CCTTAACCCCATCTCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCA	1171
Db	2521	CCTTAACCCCATCTCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCAGCAATACCA	2580
Qy	1172	TAAGAATTACTAAAAAATAAAAAAAAAA	1197
Db	2581	TAAAGATTACTAAAAAATAAAAAAAAAA	2606
RESULT 6			
US-09-976-674-12			
; Sequence 12, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 12			
; LENGTH: 4829			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-976-674-12			
Query Match 86.3%; Score 1033.2; DB 3; Length 4829;			
Best Local Similarity 88.7%; Pred. No: 0;			
Matches 1194; Conservative 0; Mismatches 3; Indels 149; Gaps 1;			
Qy	1	ATTTTGAAGGCACCAAGACTCCCTTTTGAAGCATCCTCTGTAGCATCCTCTGTAGTGTAGTACGTAA	60
Db	1781	ATTTTGAAGGCACCAAGACTCCCTTTTGAAGCATCCTCTGTAGCATCCTCTGTAGTGTAGTACGTAA	1840
Qy	61	ATCTCTGGAGGTGACAAAGGCTGACTGAGCGGTGCTACATCTTCTGTGTCATCAGTC	120
Db	1841	ATCTCTGGAGGTGACAAAGGCTGACTGAGCGGTGCTACATCTTCTGTGTCATCAGTC	1900
Qy	121	AGCACTGTGACTTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTCTCC	180
Db	1901	AGCACTGTGACTTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTCTCC	1960
Qy	181	TTTAAAGCTATCAAGTCTCTGAAAGTGAACCAACTTGCACAAAGAAATTTTGGGCA	240
Db	1961	TTTAAAGCTATCAAGTCTCTGAAAGTGAACCAACTTGCACAAAGAAATTTTGGGCA	2020
Qy	241	GAATTTTGAATTCAGCAGGTCTCTTCTGTGACTATACTCTCCAGAAATTTTCTTTTG	300
Db	2021	CCATTTTGAATTCAGCAGGTCTCTTCTGTGACTATACTCTCCAGAAATTTTCTTTTG	2080
Qy	301	AAAGTACTACTGATTTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTG	360
Db	2081	AAAGTACTACTGATTTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTG	2140

361 GAAAGAAATATCTACTGCTGTTTCATATATGTTGGTCCCTC----- 402  
Db |  
2141 GAAAGAAATATCTACTGCTGTTTCATATATGTTGGTCCCTCAGGTGAGTTGGTGA 2200  
Qy |  
403 ----- 402  
Db |  
2201 TAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAATACCTAGCCTCTCTAGGTTATGT 2260  
Qy |  
403 ----- 402  
Db |  
2261 GGTGTAGTATAGACACAGAGGGATCCGTGTCACGAGGGCTTAAATTTGAAGCGCCTT 2320  
Qy |  
403 -----AGGTTCAAATAGAAATGACGATCAGGTGGAAGACTCCAATATCTPAGC 451  
Db |  
2321 TAAATATAAATGGGTCAAATAGAAATGACGATCAGGTGGAAGACTCCAATATCTPAGC 2380  
Qy |  
452 TTTCTGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCCCTATGGAGG 511  
Db |  
2381 TTTCTGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCCCTATGGAGG 511  
Qy |  
512 ATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGG 571  
Db |  
2441 ATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGG 2500  
Qy |  
572 GGCCTCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACTTTATATGGGTCA 631  
Db |  
2501 GGCCTCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACTTTATATGGGTCA 2560  
Qy |  
632 CCTGTACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTT 691  
Db |  
2561 CCTGTACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTT 2620  
Qy |  
692 CCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGAGGTGGAAGCCATATGATTTA 751  
Db |  
2621 CCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGAGGTGGAAGCCATATGATTTA 2680  
Qy |  
752 TGCATATACAGTATATTTACTGAGTTTCTTCTAGGCTGGAAGCCATATGATTTA 811  
Db |  
2681 TGCATATACAGTATATTTACTGAGTTTCTTCTAGGCTGGAAGCCATATGATTTA 2740  
Qy |  
812 GATCTATCTCAGGAGAGACAGATGAAGTTTCTGATTCGGAGAACATTTATGAAT 871  
Db |  
2741 GATCTATCTCAGGAGAGACAGATGAAGTTTCTGATTCGGAGAACATTTATGAAT 2800  
Qy |  
872 GATCTTTTGGCTACCTTTCAAGAAAACCTTTGGATCAGTATTTGCTGCTCTAAAGTGAT 931  
Db |  
2801 GATCTTTTGGCTACCTTTCAAGAAAACCTTTGGATCAGTATTTGCTGCTCTAAAGTGAT 2860  
Qy |  
932 ATAAATTTGCACTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAAAATGAGGAGG 991  
Db |  
2861 ATAAATTTGCACTGTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAAAATGAGGAGG 2920  
Qy |  
992 TTTAATCAACAGAAAACAGAAATTTGATCATCATTTTGGATGATTAATGATTAATCT 1051  
Db |  
2921 TTTAATCAACAGAAAACAGAAATTTGATCATCATTTTGGATGATTAATGATTAATCT 2980  
Qy |  
1052 ACTCTGMAAAATAAATGTGGTCCATGAGGGGTCTACGGTTTGTGGTATTAATCTAATA 1111  
Db |  
2981 ACTCTGMAAAATAAATGTGGTCCATGAGGGGTCTACGGTTTGTGGTATTAATCTAATA 3040  
Qy |  
1112 CCTTAAACCCCATCTCTCAAAATCAATATGATACATATTTCTGAGAGCCAGCAATACCA 1171  
Db |  
3041 CCTTAAACCCCATCTCTCAAAATCAATATGATACATATTTCTGAGAGCCAGCAATACCA 3100  
Qy |  
1172 TAAGATTAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 1197  
Db |  
3101 TAAGATTAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 3126

RESULT 7  
US-09-976-674-22

; Sequence 22, Application US/09976674

; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR APPLICATION NUMBER: 2001-10-12  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 4685  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-22

Query Match 84.7%; Score 1013.6; DB 3; Length 4685;  
Best Local Similarity 91.3%; Pred. No. 0;  
Matches 1098; Conservative 0; Mismatches 99; Indels 5; Gaps 2;  
Qy 1 ATTTGAAGGCACCAAGAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAA 60  
Db 1781 ATTTGAAGGCACCAAGAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAA 1840  
Qy 61 ATCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATCTGTGTGATCAGTC 120  
Db 1841 ATCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATCTGTGTGATCAGTC 1900  
Qy 121 AGCACTGTACATCTTTTATAGTATGTAATGTAACAGAGAAATCCACATCTGTGTGCC 180  
Db 1901 AGCACTGTACATCTTTTATAGTATGTAATGTAACAGAGAAATCCACATCTGTGTGCC 1960  
Qy 181 TTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTGTGCAAAACAAAGAAATTTTGGGCCA 240  
Db 1961 TTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTGTGCAAAACAAAGAAATTTTGGGCCA 2020  
Qy 241 CCATTTTGGATTCAG----CAGGTCTCTCTCTCTGACT-ATACTCTCTCAGAAATTTTCTC 295  
Db 2021 CCATTTTGGATTCAGTCT 2080  
Qy 296 TTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTTCAAGGCTCATGATCTACA 355  
Db 2081 TTTTCCGTTTGAATACCTAGCCCTCTCTAGTTTATGTTGTTAGTGATAGACACAGGGG 2140  
Qy 356 GCCTGAAAGAAATATCTCTACTGTCTCTTTCATATATGTTGGTCTCTCAGGGTCAAAATAGA 415  
Db 2141 ATCTCTCACCGAGGCTTAAATTTGAAGGCGCCTTTAAATATAAAATGGGTCAAAATAGA 2200  
Qy 416 AATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCTTGAAT 475  
Db 2201 AATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCTTGAAT 2260  
Qy 476 AGATCTGTGGGCATCCAGGCTGGTCTTATGGAGTACCTCTCCCTGATGGCATTAAT 535  
Db 2261 AGATCTGTGGGCATCCAGGCTGGTCTTATGGAGTACCTCTCCCTGATGGCATTAAT 2320  
Qy 536 GCAGAGGTGAGATATCTTCAAGGTTGCTTATGCTGGGGCCCCAGTCACTCTGTGGATCTT 595  
Db 2321 GCAGAGGTGAGATATCTTCAAGGTTGCTTATGCTGGGGCCCCAGTCACTCTGTGGATCTT 2380  
Qy 596 CTATGATACAGGATACAGGAACTTATATGGGTCACTCTGACCAAGATGAACAGGGCTA 655  
Db 2381 CTATGATACAGGATACAGGAACTTATATGGGTCACTCTGACCAAGATGAACAGGGCTA 2440  
Qy 656 TTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTTCCCTCTGAAACAAATCGTTTACT 715  
Db 2441 TTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTTCCCTCTGAAACAAATCGTTTACT 2500  
Qy 716 GCTCTTACATGTTTCTCTGGATGAGAAATGTCCTATTTTGCATACCATCATATATCTAGT 775



Db 2501 GCTTACATGGTTTCTGGATGGAATGTCATTTTGACATACCAGTATATTACTGAG 2560  
Qy 776 TTTTATAGTGGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGACACAG 835  
Db 2561 TTTTATAGTGGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGACACAG 2620  
Qy 836 CATAGAGTTCTTGAATCGGAGAACATTTATGAACTGTCATCTTTTGCACTACCTTCAAGA 895  
Db 2621 CATAGAGTTCTTGAATCGGAGAACATTTATGAACTGTCATCTTTTGCACTACCTTCAAGA 2680  
Qy 896 AAACCTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAATC 955  
Db 2681 AAACCTTGGATCAGCTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAATC 2740  
Qy 956 TCTGGTATACATCGGCTATTATTAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAT 1015  
Db 2741 TCTGGTATACATCGGCTATTATTAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAT 2800  
Qy 1016 TGAATCATCAATTTGATACCTGCGATGTAAACATCTACTCTCTGAAATAAATGTGGTGC 1075  
Db 2801 TGAATCATCAATTTGATACCTGCGATGTAAACATCTACTCTCTGAAATAAATGTGGTGC 2860  
Qy 1076 ATGAGGGGTCTACGGTTGTGTAGTAAATCTAATACCTTAACCCCATGCTCAAAATC 1135  
Db 2861 ATGAGGGGTCTACGGTTGTGTAGTAAATCTAATACCTTAACCCCATGCTCAAAATC 2920  
Qy 1136 AAATGATACATATCTCTGAGAGACCCAGCAATACCAATGAAGAAATTAATAAATAAATAA 1195  
Db 2921 AAATGATACATATCTCTGAGAGACCCAGCAATACCAATGAAGAAATTAATAAATAAATAA 2980  
Qy 1196 AA 1197  
Db 2981 AA 2982

## RESULT 8

US-10-070-464-6  
; Sequence 6, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1669  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-070-464-6

Query Match 73.6%; Score 881; DB 3; Length 1669;  
Best Local Similarity 87.2%; Pred. No. 1.3e-277;  
Matches 1044; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

Qy 1 ATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAAGTACCTAA 60  
Db 618 ATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAAGTACCTAA 677  
Qy 61 ATCTTGAGAGGTGACAAAGGCTGACTGACCGTGGTACTCATTCTTGTGTCATCAGTC 120  
Db 678 ATCTTGAGAGGTGACAAAGGCTGACTGACCGTGGTACTCATTCTTGTGTCATCAGTC 737

Qy 121 AGCACTGTGACTCTCTTTATAGTAAGTAAACAGAGAAATCCACACTGTGTGTC 180  
Db 738 AGCACTGTGACTCTCTTTATAGTAAGTAAACAGAGAAATCCACACTGTGTGTC 797  
Qy 181 TTTTCAAGCTATCAAGTCTCTGAAAGTACCCCACTTTGCAAAAACAAAGGAATTTTGGGCCA 240  
Db 798 TTTTCAAGCTATCAAGTCTCTGAAAGTACCCCACTTTGCAAAAACAAAGGAATTTTGGGCCA 857  
Qy 241 CCAATTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTG 300  
Db 858 CCAATTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTTG 917  
Qy 301 AAACTACTACTGGATTTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTACAGCCTG 360  
Db 918 AAACTACTACTGGATTTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTACAGCCTG 977  
Qy 361 GAAAGAAATATCTTACTGTCTGCTTATATATATGCTGCTCTCAGGGTCAATATAGAAATG 420  
Db 978 GAAAGAAATATCTTACTGTCTGCTTATATATATGCTGCTCTCA----- 1020  
Qy 421 ACGATCAGGTGGAAGGACTCCAATATCTAGTCTTCTCGATATGATTTTCATTGACTTAGATC 480  
Db 1021 ----- 1020  
Qy 481 GTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAATTAATGCAGA 540  
Db 1021 ----- 1020  
Qy 541 GGTCAAGATATCTTCAGGGTTGCTATTGTGGGGCCCGAGTCACTCTGTGGATCTTCTATG 600  
Db 1021 -----GGTTGCTATTGTCTGGGGCCCGAGTCACTCTGTGGATCTTCTATG 1064  
Qy 601 ATACAGGATACACGGAACGTTATATGGGTCACTCCTGACCAAGTAAAGAGGCTATTACT 660  
Db 1065 ATACAGGATACACGGAACGTTATATGGGTCACTCCTGACCAAGTAAAGAGGCTATTACT 1124  
Qy 661 TAGGATCTGTGGCCATGCAAGAGCAAGAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 720  
Db 1125 TAGGATCTGTGGCCATGCAAGAGCAAGAGTTCCCTCTGAAACCAATCGTTTACTGCTCT 1184  
Qy 721 TACATGTTTCTCGATGAGAAATGTCATTTTGCACATACACAGTATATTACTGAGTTTT 780  
Db 1185 TACATGTTTCTCGATGAGAAATGTCATTTTGCACATACACAGTATATTACTGAGTTTT 1244  
Qy 781 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAA 840  
Db 1245 TAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAA 1304  
Qy 841 GAGTTCCTGAATCGGAGAACATTTATGAATCTGATCTTTTGCACTTACCTTCAAGAAACC 900  
Db 1305 GAGTTCCTGAATCGGAGAACATTTATGAATCTGATCTTTTGCACTTACCTTCAAGAAACC 1364  
Qy 901 TTGGATCACTGTTGCTGCTTAAAGTATATAATTTTGAAGTGTAGTAACTCTCTGG 960  
Db 1365 TTGGATCACTGTTGCTGCTTAAAGTATATAATTTTGAAGTGTAGTAACTCTCTGG 1424  
Qy 961 TATACACTGGCTATTTAACCAAAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATC 1020  
Db 1425 TATACACTGGCTATTTAACCAAAATGAGGAGTTTAAATCAACAGAAAACACAGAAATGATC 1484  
Qy 1021 ATCACTTTTGATACCTGCGCATGTAACTACTCTCTGAAATAAATATGTTGGTGCATGCA 1080  
Db 1485 ATCACTTTTGATACCTGCGCATGTAACTACTCTCTGAAATAAATATGTTGGTGCATGCA 1544  
Qy 1081 GGGGTCTACGGTTTGTGGTGTAGTAACTTAACTTAAACCCACATGCTCAAAATCAATG 1140  
Db 1545 GGGGTCTACGGTTTGTGGTGTAGTAACTTAACTTAAACCCACATGCTCAAAATCAATG 1604  
Qy 1141 ATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAAATAAATAAATAA 1197  
Db 1605 ATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAAAATAAATAAATAA 1661



RESULT 9									
US-09-976-674-8									
; Sequence 8, Application US/09976674									
; Patent No. 6844180									
; GENERAL INFORMATION:									
; APPLICANT: Qi, Steve									
; APPLICANT: Akinsanya, Karen									
; APPLICANT: Riviere, Pierre									
; APPLICANT: Junien, Jean-Louis									
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV									
; FILE REFERENCE: 70669									
; CURRENT APPLICATION NUMBER: US/09/976,674									
; PRIOR FILING DATE: 2001-10-12									
; PRIOR APPLICATION NUMBER: US 60/240,117									
; NUMBER OF SEQ ID NOS: 61									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 8									
; LENGTH: 4523									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-976-674-8									
Query Match 72.9%; Score 873; DB 3; Length 4523;									
Best Local Similarity 86.3%; Pred. No. 1e-274;									
Matches 1040; Conservative 0; Mismatches 0; Indels 157; Gaps 1;									
Qy	1	ATTTGAGGCCCAAGACTCCCTTTAGAGCATCACCCTGCTAGCTAGTCAAGTACGTAA	60						
Db	1781	ATTTGAGGCCCAAGACTCCCTTTAGAGCATCACCCTGCTAGCTAGTCAAGTACGTAA	1840						
Qy	61	ATCTGGAGGTTGACAGGCTGACGACCGTGGTACTCAATCTTGTGTGATCAGTC	120						
Db	1841	ATCTGGAGGTTGACAGGCTGACGACCGTGGTACTCAATCTTGTGTGATCAGTC	1900						
Qy	121	AGCACTGTGACTTCTTTAAGTATAGTATAGTAAACAGAGAAATCCACATGTGTGCC	180						
Db	1901	AGCACTGTGACTTCTTTAAGTATAGTATAGTAAACAGAGAAATCCACATGTGTGCC	1960						
Qy	181	TTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAAAGAAATTTTGGGCCA	240						
Db	1961	TTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAAAGAAATTTTGGGCCA	2020						
Qy	241	CCATTTTGGATTCCAGAGTCTCTCTCTGACTATCTCTCTGACTATCTCTCTCTCTCTCTCT	300						
Db	2021	CCATTTTGGATTCCAGAGTCTCTCTCTGACTATCTCTCTGACTATCTCTCTCTCTCTCTCT	2080						
Qy	301	AAAGTACTACTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG	360						
Db	2081	AAAGTACTACTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTG	2140						
Qy	361	GAAGAAATATCTACTGCTGCTTTCATATATATGTTGTTCTCAGGGTCAATAGAAATTG	420						
Db	2141	GAAGAAATATCTACTGCTGCTTTCATATATATGTTGTTCTCAGGGTCAATAGAAATTG	2179						
Qy	421	ACGATCAGTGGNAGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATC	480						
Db	2180	-----	2179						
Qy	481	GTGTGGGCATCCACGGCTGGTCTATGAGGAGTACCTCTCCTGATGGCAATTAATGAGA	540						
Db	2180	-----	2179						
Qy	541	GGTCAGATATCTTCAGGGTGTCTATTGCTGGGGGCCAGTCACTCTGTGATCTTCTATG	600						
Db	2180	-----GGTGTGATTTGCTGGGGGCCAGTCACTCTGTGATCTTCTATG	2223						
Qy	601	ATACAGGATACACGGAACTTATATGGGTACCCCTGACCAAGTGAACAGGGGTATTACT	660						
Db	2224	ATACAGGATACACGGAACTTATATGGGTACCCCTGACCAAGTGAACAGGGGTATTACT	2283						
Qy	661	TAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGTAACCAAAATCGTTTACTGTCT	720						
Db		-----							

Db	2284	TAGGATCTGTGGCCATGCAAGCAGAAAAAGTTC	2343
Qy	721	TACATGGTTTCTGGATGAGAAATGTCATTTTGCACATACAGTATATATCTAGTGTGTTT	780
Db	2344	TACATGGTTTCTGGATGAGAAATGTCATTTTGCACATACAGTATATATCTAGTGTGTTT	2403
Qy	781	TAGTGGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAA	840
Db	2404	TAGTGGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAA	2463
Qy	841	GAGTTCCTGAATCGGAGAACATTATGAATCGATCTTTTGCACTACCTTCAAGAAAAACC	900
Db	2464	GAGTTCCTGAATCGGAGAACATTATGAATCGATCTTTTGCACTACCTTCAAGAAAAACC	2523
Qy	901	TTGGATCACGTATTTGCTCTTAAAGTGTATATAATTTTGACCTGTGTAGAACTCTCTGG	960
Db	2524	TTGGATCACGTATTTGCTCTTAAAGTGTATATAATTTTGACCTGTGTAGAACTCTCTGG	2583
Qy	961	TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATTCATC	1020
Db	2584	TATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATTCATC	2643
Qy	1021	ATCACATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAATGTGTGCCATGCA	1080
Db	2644	ATCACATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAATGTGTGCCATGCA	2703
Qy	1081	GGGGTCTACGGTTTGGTGTAGTAATCTTAATCCTTTAAACCCACATGCTCAAAATCAAATG	1140
Db	2704	GGGGTCTACGGTTTGGTGTAGTAATCTTAATCCTTTAAACCCACATGCTCAAAATCAAATG	2763
Qy	1141	ATACATATTTCTGAGAGACCCAGCAATACCAATAGAAATTTCTAAAAAATAAAAAA	1197
Db	2764	ATACATATTTCTGAGAGACCCAGCAATACCAATAGAAATTTCTAAAAAATAAAAAA	2820
RESULT 10			
US-09-976-674-2			
; Sequence 2, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 2			
; LENGTH: 2671			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-976-674-2			
Query Match 64.9%; Score 776.4; DB 3; Length 2671;			
Best Local Similarity 86.3%; Pred. No. 3.3e-243;			
Matches 934; Conservative 0; Mismatches 1; Indels 147; Gaps 1;			
Qy	1	ATTTGAGGCCCAAGACTCCCTTTAGAGCATCACCCTGCTAGCTAGTCAAGTACGTAA	60
Db	1575	ATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTAGCTAGTCAAGTACGTAA	1634
Qy	61	ATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTC	120
Db	1635	ATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTC	1694
Qy	121	AGCACTGTGACTTCTTTTATAAGTATAGTATAGTAAACAGAAATCCACATGTGTGTCCC	180
Db	1695	AGCACTGTGACTTCTTTTATAAGTATAGTATAGTAAACAGAAATCCACATGTGTGTCCC	1754

```
QY 181 TTACAACTATCAAGTCTGAAGTACCACTGCAAAACAAAGGAATTTGGGCA 240
|
|
|
Db 1755 TTACAACTATCAAGTCTGAAGTACCACTGCAAAACAAAGGAATTTGGGCA 1814
|
|
|
QY 241 CCATTTGGATTCAGCAGGCTCTCTCTGATATATCTCTCCAGAAAATTTCTCTTTG 300
|
|
|
Db 1815 CCATTTGGATTCAGCAGGCTCTCTCTGATATATCTCTCCAGAAAATTTCTCTTTG 1874
|
|
|
QY 301 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTG 360
|
|
|
Db 1875 AAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTG 1934
|
|
|
QY 361 GAAAGAAATATCTACTGCTGCTTTCATATATGCTGCTCA----- 403
|
|
|
Db 1935 GAAAGAAATATCTACTGCTGCTTTCATATATGCTGCTCAAGGTGGTGAATA 1994
|
|
|
QY 404 ----- 403
|
|
|
Db 1995 ATCGATTTAAGGAGTCAAGTATTTCCGCTTGGAATACCTAGCTCTCTAGGTTATGTGG 2054
|
|
|
QY 404 ----- 403
|
|
|
Db 2055 TTGTAGTAGATAGACACAGGGATCTGTACCGAGGGCTTAAATTTGAAGCGGCTTTA 2114
|
|
|
QY 404 -----GGGTCAAATAGAAATTGAAGATCAGGTGGAAGGATCCAAATATCTAGCTT 453
|
|
|
Db 2115 AATATAAAATGGGTCAAATAGAAATTGACGATCAGGTGGAAGGATCCAAATATCTAGCTT 2174
|
|
|
QY 454 CTGATATGATTTCAATTCATTTAGATCGTGTGGGATCCAGGCTGGTCTTATGAGGAT 513
|
|
|
Db 2175 CTGATATGATTTCAATTCATTTAGATCGTGTGGGATCCAGGCTGGTCTTATGAGGAT 2234
|
|
|
QY 514 ACCTCTCCCTGATGGATTAATGCAGAGTTCAGATATCTTCAGGCTGCTATTCCTGGG 573
|
|
|
Db 2235 ACCTCTCCCTGATGGATTAATGCAGAGTTCAGATATCTTCAGGCTGCTATTCCTGGG 2294
|
|
|
QY 574 CCCAGTCACCTCTGTGGATCTCTATGATACAGGATACACGGAAGTTATATGGTCAAC 633
|
|
|
Db 2295 CCCAGTCACCTCTGTGGATCTCTATGATACAGGATACACGGAAGTTATATGGTCAAC 2354
|
|
|
QY 634 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCTCAAGCAGAAAAGTTCC 693
|
|
|
Db 2355 CTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCTCAAGCAGAAAAGTTCC 2414
|
|
|
QY 694 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGATGAGAAATGTCATTTG 753
|
|
|
Db 2415 CCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGATGAGAAATGTCATTTG 2474
|
|
|
QY 754 CACATACAGTATATTACTGAGTTTTTATGAGGCTGGAAGCCATGATTTACAGA 813
|
|
|
Db 2475 CACATACAGTATATTACTGAGTTTTTATGAGGCTGGAAGCCATGATTTACAGA 2534
|
|
|
QY 814 TCTATCTCAGGAGACAGCAGTAAAGATTTCTGATCGGGGAACATTTAGACTGC 873
|
|
|
Db 2535 TCTATCTCAGGAGACAGCAGTAAAGATTTCTGATCGGGGAACATTTAGACTGC 2594
|
|
|
QY 874 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTCCTGCTCTAAAAGTGATAT 933
|
|
|
Db 2595 ATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTCCTGCTCTAAAAGTGATAT 2654
|
|
|
QY 934 AA 935
|
|
|
Db 2655 GA 2656
```

RESULT 11  
US-09-280-116-171  
; Sequence 171, Application US/09280116A  
; Patent No. 6331427  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs

FILE REFERENCE: 5800-24, 035800/176965  
; CURRENT APPLICATION NUMBER: US/09/280,116A  
; CURRENT FILING DATE: 1999-03-26  
; NUMBER OF SEQ ID NOS: 268  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 171  
; LENGTH: 823  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: prolyl oligopeptidases  
US-09-280-116-171

Query Match 52.9%; Score 633.4; DB 3; Length 823;  
Best Local Similarity 94.3%; Pred. No. 1.1e-196;  
Matches 776; Conservative 0; Mismatches 11; Indels 36; Gaps 10;

QY 406 GTCAAAATAGAANTTACGATCAGGTGGAAGGAC-TCCAATATCTAGCTTCTCGATATGAT 464  
|  
|  
|  
Db 1 GTCAAAATAGAANTTACGATCAGGTGGAAGGACATCCAATATCTAGCTTCTCGATATGAT 60  
|  
|  
|  
QY 465 TTCAATTGACTTATGATCGTGTGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTG 524  
|  
|  
|  
Db 61 TTCAATTGACTTATGATCGTGTGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTG 120  
|  
|  
|  
QY 525 ATGGCATTAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACT 584  
|  
|  
|  
Db 121 ATGGCATTAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGTCCCACTCACT 180  
|  
|  
|  
QY 585 CTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCTGACAGAAAT 644  
|  
|  
|  
Db 181 CTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCTGACAGAAAT 240  
|  
|  
|  
QY 645 GAAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGAACCA 704  
|  
|  
|  
Db 241 GAAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGAACCA 300  
|  
|  
|  
QY 705 AATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTTGCACATACCACT 764  
|  
|  
|  
Db 301 AATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCATTTTGCACATACCACT 360  
|  
|  
|  
QY 765 ATATTACTAGTTTTTTAGTGGGCTGGAAGCCATATGATTTTAC----- 810  
|  
|  
|  
Db 361 ATATTACTAGTTTTTTAGTGGGCTGGAAGCCATATGATTTTACAGTATCTTTATTTT 420  
|  
|  
|  
QY 811 -----AGATCTATCTCTCA-GGAGAGACA-CAGCATAGAGTTTCTG--AATCG 854  
|  
|  
|  
Db 421 TTGTTTGTGTTGTAAGATCTATCTCTCAGGAGAGACACCGGCATAAGAGTTCCCTGAATCG 480  
|  
|  
|  
QY 855 GGAGAACATTTATGAACCTGCACTTTT-GACTACCTTCAAG-AAACCTTGGATCA-CGT 911  
|  
|  
|  
Db 481 GGAGAACATTTATGAACCTGCACTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACCTT 540  
|  
|  
|  
QY 912 ATTGCTGCT-CTAAAAGTGATATAAATTTGACCTGCTGTAGAACTCTCT-GGTATACACTG 969  
|  
|  
|  
Db 541 ATTGCTGCTCTAAAAGTGATATAAATTTGACCTGCTGTAGAACTCTCTGGGGTATACACTG 600  
|  
|  
|  
QY 970 GCTATTAAACCAATAGAGGAGTTTAAATCAACAGAAAAACAGAGAATTGATCATCACTTT 1029  
|  
|  
|  
Db 601 GCTATTAAACCAATAGAGGAGTTTAAATCAACAGAAAAACAGAGAATTTGATCATCACTTT 660  
|  
|  
|  
QY 1030 TGATACCTGCTGCTAAACATCTACTCTGAAAAATAAATGTTGGTGCATGCAAGGGGCTTAC 1089  
|  
|  
|  
Db 661 TGATACCTGCTGCTAAACATCTACTCTGAAAAATAAATGTTGGTGCATGCAAGGGGCTTAC 720  
|  
|  
|  
QY 1090 GGTTTGGTGTAGTAATCTAAATACCTTAAGCCCACTGCTCAAAAATCAAAATGATACATATT 1149  
|  
|  
|  
Db 721 GGTTTGGTGTAGTAATCTAAATACCTTAAGCCCACTGCTCAAAAATCAAAATGATACATATT 780  
|  
|  
|  
QY 1150 CTTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAAATAA 1192  
|  
|  
|  
Db 781 CTTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAAATAA 823

```
RESULT 12
US-10-070-464-8
; Sequence 8, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-8

Query Match      38.3%; Score 459; DB 3; Length 1083;
Best Local Similarity 80.7%; Pred. No. 1.8e-139;
Matches 616; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 60
DB 321 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAA 380
QY 61 ATCTGTGAGAGGTGCAAGAGCTGACTGACCGTGGTACTCACTTCTTGTGTCATCAGTC 120
DB 381 ATCTGTGAGAGGTGCAAGAGCTGACTGACCGTGGTACTCACTTCTTGTGTCATCAGTC 440
QY 121 AGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCC 180
DB 441 AGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCC 500
QY 181 TTTTACAAGCTATCAAGTCTGGAAGATGACCAACTTGGAAACAAAGAAATTTTGGGCCA 240
DB 501 TTTTACAAGCTATCAAGTCTGGAAGATGACCAACTTGGAAACAAAGAAATTTTGGGCCA 560
QY 241 CCATTTTGGATTACAGAGGTCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTG 300
DB 561 CCATTTTGGATTACAGAGGTCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTG 620
QY 301 AAGTACTAGTGGATTACATTTGTATGGATGCTCTACAAGCTCATGATCAGAGCTG 360
DB 621 AAGTACTAGTGGATTACATTTGTATGGATGCTCTACAAGCTCATGATCAGAGCTG 680
QY 361 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTGGTCTCA----- 403
DB 681 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTGGTCTCA----- 740
QY 404 ----- 403
DB 741 ATCGGTTTAAAGGAGTCAAGATATTCGCTTGAATACCTAGACCTCTCTAGGTTATGTGG 800
QY 404 ----- 403
DB 801 TTGTAGTGATAGACACAGGGGATCCTGTCAACCGAGGCTTAAATTTTGAAGGGCCCTTTA 860
QY 404 -----GGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTT 453
DB 861 AATATAAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTT 920
QY 454 CTCGATATGATTTCAITGACTTAGATCGTGTGGGATCCAGGGTGGTCTCTATGGAGGAT 513
DB 921 CTCGATATGATTTCAITGACTTAGATCGTGTGGGATCCAGGGTGGTCTCTATGGAGGAT 980
```

```
QY 514 ACCTCTCCCTGATGGCATTAATGACAGAGGTGAGATATCTTCAGGGTGTCTATTCTCGGG 573
DB 981 ACCTCTCCCTGATGGCATTAATGACAGAGGTGAGATATCTTCAGGGTGTCTATTCTCGGG 1040
QY 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 616
DB 1041 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083

RESULT 13
US-09-392-184-31/c
; Sequence 31, Application US/09392184
; Patent No. 6395889
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: PROTEASE HOMOLOGS
; FILE REFERENCE: 5800-55
; CURRENT APPLICATION NUMBER: US/09/392,184
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(612)
; OTHER INFORMATION: prolyloloigo (prolyl oligopeptidase)
; NAME/KEY: misc feature
; LOCATION: (1)...(612)
; OTHER INFORMATION: n = A,T,C or G
US-09-392-184-31

Query Match      27.3%; Score 326.8; DB 3; Length 612;
Best Local Similarity 79.2%; Pred. No. 2.6e-96;
Matches 397; Conservative 0; Mismatches 101; Indels 3; Gaps 3;

QY 413 AGAAATTTGACGATCAGGTGGAAGGACT-CCAAATATCTAGCTTCTCGATATGATTTCA 471
DB 520 AGAATTCGCGATCAGGTGCGGAGTCCCAAATATCTAACTTCTCGAAATGATTTCA 461
QY 472 ACTTAGATCTGTGGGCTCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCAT 531
DB 460 ACNTAAGATCTGTGGCAATCCNCGCTGTCTCTATGGAGGATACCTCTCCCTGATGGCAN 401
QY 532 TAATGACAGAGTTCAGATATCTTCAGGGTGTCTATGCTGGGCCCCAGTCACTCTGTGGA 591
DB 400 TAATTCAGAGTTCAGATATCTTCAGGGT-CTAATTTCTGGGCCCCAGTCACTCTGTGGA 342
QY 592 TCTTCTATGATACAGGATACAGGACGTTATATGGTCACTGACCCAGGATGAAACAGG 651
DB 341 TCTTCTATGAAACAGGAAACNCGGACGTTANATGGGTCACTTNNCCAGNATGGACAGG 282
QY 652 GCTATTTACTTAGATCTGTGGCATGCGAGGAAAGTTCCTCTGAAACCAATCGTT 711
DB 281 GCTATTTACTTAGATCTGTGGCATTCANAGGAAAGTTCCTCTGNNCCCAATNNGT 222
QY 712 TACTGCTCTTACATGGTTCCTGGATGAGATGTCATTTTGCAATACCAATATATAC 771
DB 221 TNCCTCTCTANATGGTTCCTGGATTAGATTTCCANATTTTCANACCAAGNANATAC 162
QY 772 TGAGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGATCATCTCAGGAGAC 831
DB 161 TNAGGTTTTTTAGTGAGGGCTGGAAGCCAAATGANTNACAGAACCACTCNGAGAGN 102
QY 832 ACAGCATAAAGAGTTCCTGAATCGGGAGAAACATTATG-AACTGCAATCTTTTGACTACCTT 890
DB 101 CCAACANNAGAGTTCCTGAATCGGGAGGACAAATATNTGACCTNNACCTTTNACNACCN 42
QY 891 CAAGAAAAACCTTGGATCAGCT 911
```

Db 41 CAGAAACCTTGGATCAGT 21  
|||||

## RESULT 14

US-09-621-976-1787  
; Sequence 1787, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET 054PE2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 1787  
; LENGTH: 308  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 17..178  
; NAME/KEY: sig\_peptide  
; LOCATION: 17..133  
; OTHER INFORMATION: Von Heljne matrix  
; OTHER INFORMATION: score 4.59999990463257  
; OTHER INFORMATION: seq AHTSILSLFLVRA/GK  
US-09-621-976-1787

Query Match 24.8%; Score 297; DB 3; Length 308;  
Best Local Similarity 99.7%; Pred. No. 9.9e-87;  
Matches 308; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 659 CTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCT 718  
|||||  
Db 1 CTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCT 60  
|||||  
QY 719 CTTACATGGTTTCTCGATGAGAAATGTCATTTGGACATACCAAGTATTTACTGAGTTT 778  
|||||  
Db 61 CTTACATGGTTTCTCGATGAGAAATGTCATTTGGACATACCAAGTATTTACTGAGTTT 120  
|||||  
QY 779 TTTAGTGAAGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCAT 838  
|||||  
Db 121 TTTAGTGAAGGCTGGAAG- CATATGATTTACAGATCTATCTCAGGAGAGACACAGCAT 179  
|||||  
QY 839 AAGAGTTCTCGAATCGGAGAACATTATGACTGATCTTTTGGCACTACCTTCAAGAAA 898  
|||||  
Db 180 AAGAGTTCTCGAATCGGAGAACATTATGACTGATCTTTTGGCACTACCTTCAAGAAA 239  
|||||  
QY 899 CTTTGGATCAGTATCTCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 958  
|||||  
Db 240 CTTTGGATCAGTATCTCTCTCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCT 299  
|||||  
QY 959 GGTATACAC 967  
|||||  
Db 300 GGTATACAC 308  
|||||

## RESULT 15

US-09-976-674-4  
; Sequence 4, Application US/09976674  
; Patent No. 684180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674

; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117.  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2617  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-4

Query Match 17.8%; Score 212.8; DB 3; Length 2617;  
Best Local Similarity 63.9%; Pred. No. 1.4e-58;  
Matches 322; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 404 GGGTCAAAATAGAAATTCAGATCAGGTGGAGAGACTCCCAATATCTAGTCTTCGTGATAGA 463  
|||||  
Db 2101 GGGCCAGGTGGAGATCGAGGACCGAGGTGGAGGGCTTGCAGTTCTGTTGGCCGAGAGATGG 2160  
|||||  
QY 464 TTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTCGTCTATGGAGGATACCTCTCCCT 523  
|||||  
Db 2161 CTTTCATCGACCTGAGCGGAGTTGCCATCCATGGCTGTCTACGGGGGCTTCTCTCGCT 2220  
|||||  
QY 524 GATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCCAC 583  
|||||  
Db 2221 CATGGGGCTAATCCCAAGCCCCCAGGTTTCAAGGTGGCCATCGCGGGGTGCCCCCGGTCCAC 2280  
|||||  
QY 584 TCTGTGGATCTTCTATGATACAGGATACAGGAACTGATATATGGGTACCCCTGACACAGAA 643  
|||||  
Db 2281 CGTCTGGATGGCCTTACGACACAGGTACACTGAGCCCTACATGGAGCTCCCTTGAGAACAA 2340  
|||||  
QY 644 TGAACAGGGCTATTATTAGGATCTGTGGCCATGTCAGGAGCAAGCAAGAAAAGTTTCCCTCTGAACC 703  
|||||  
Db 2341 CCAGCAGGCTATGAGCGGGTTTCOGTGGCCCTGCACGTGGAGAGAGTGGCCCAATGAGCC 2400  
|||||  
QY 704 AAATCGTTTACTGCTCTTACATGCTTTCCTGGATGAGAAATGCCATTTTGGACATACCAG 763  
|||||  
Db 2401 CAACCGCTTGTCTTATCTCCACGGCTTCTTGGACGAAAACGTGCACTTTTCCACACAAA 2460  
|||||  
QY 764 TATATTACTGAGTTTTTTAGTGAAGGCTGGAAGCCATATGATTTACAGATCTATCTCA 823  
|||||  
Db 2461 CTTCTCTCTCCCACTGATCCGAGCAGGAAACCTTACCAGCTCCAGATCTACCCCAA 2520  
|||||  
QY 824 GGAGAGACACAGCATAGAGTTCTCTGAATCGGGAGAACATTATGAAGTGCATCTTTTGA 883  
|||||  
Db 2521 CGAGAGACACAGTATTCGCTGCCCCGAGTCGGCGGAGCACTATGAAGTCAAGTCTGCTGCA 2580  
|||||  
QY 884 CTACCTTCAAGAAAACCTTGGATC 907  
|||||  
Db 2581 CTTTCTACAGGATACCTCTGAGC 2604  
|||||

Search completed: May 3, 2006, 02:12:56  
Job time : 388.227 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:49:39 ; Search time 1635 Seconds  
(without alignments)  
2981.544 Million cell updates/sec

Title: US-10-825-632-4  
Perfect score: 1197  
Sequence: 1 attttgaaggcccaagac.....ttactaataaaaaaaaaa 1197

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA\_New.\*  
1: /SIDSS5/ptodata/2/pubpna/US08\_NEW\_PUB.seq1.\*  
2: /SIDSS5/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /SIDSS5/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /SIDSS5/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
5: /SIDSS5/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
6: /SIDSS5/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
7: /SIDSS5/ptodata/2/pubpna/US09\_NEW\_PUB.seq1.\*  
8: /SIDSS5/ptodata/2/pubpna/US09\_NEW\_PUB.seq2.\*  
9: /SIDSS5/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
10: /SIDSS5/ptodata/2/pubpna/US10\_NEW\_PUB.seq1.\*  
11: /SIDSS5/ptodata/2/pubpna/US10\_NEW\_PUB.seq2.\*  
12: /SIDSS5/ptodata/2/pubpna/US10\_NEW\_PUB.seq3.\*  
13: /SIDSS5/ptodata/2/pubpna/US10\_NEW\_PUB.seq4.\*  
14: /SIDSS5/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
15: /SIDSS5/ptodata/2/pubpna/US11\_NEW\_PUB.seq1.\*  
16: /SIDSS5/ptodata/2/pubpna/US11\_NEW\_PUB.seq2.\*  
17: /SIDSS5/ptodata/2/pubpna/US11\_NEW\_PUB.seq3.\*  
18: /SIDSS5/ptodata/2/pubpna/US11\_NEW\_PUB.seq4.\*  
19: /SIDSS5/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1040	86.9	3143	17	US-11-151-601-19 Sequence 19, Appl
2	778	65.0	2649	17	US-11-151-601-21 Sequence 21, Appl
C 3	118	9.9	1346	7	US-09-925-065A-669313 Sequence 669313, Appl
C 4	113.6	9.5	609	7	US-09-925-065A-743558 Sequence 743558, Appl
C 5	113.6	9.5	609	7	US-09-925-065A-743559 Sequence 743559, Appl
6	80.2	6.7	2238	18	US-11-079-463-1186 Sequence 1186, Ap
7	73.4	6.1	612	7	US-09-925-065A-818446 Sequence 818446, Appl
8	70	5.8	3332	18	US-11-208-288-1 Sequence 1, Appl1
9	70	5.8	3407	11	US-10-501-035-34 Sequence 34, Appl
10	68.4	5.7	2217	18	US-11-208-288-3 Sequence 3, Appl1
11	68.4	5.7	2301	10	US-10-522-789-1 Sequence 1, Appl
12	64.4	5.4	2788	9	US-10-505-928-476 Sequence 476, Appl
13	64.4	5.4	2814	17	US-11-186-284-54 Sequence 54, Appl
14	64.4	5.4	2814	18	US-11-245-147-168 Sequence 168, Appl

15	61.8	5.2	535	18	US-11-226-869-428 Sequence 428, App
16	61.6	5.1	4852	17	US-11-136-527-2130 Sequence 2130, Appl
17	61	5.1	2283	18	US-11-208-288-5 Sequence 5, Appl1
18	52.4	4.4	2778	11	US-10-932-182A-5649 Sequence 5649, Ap
19	52.4	4.4	2778	11	US-10-932-182A-5649 Sequence 5649, Ap
20	50.6	4.2	378	11	US-10-932-182A-81332 Sequence 81332, A
21	50.6	4.2	378	11	US-10-932-182A-81332 Sequence 81332, A
22	48.6	4.1	2457	11	US-10-932-182A-1107 Sequence 1107, Ap
23	48.6	4.1	2457	11	US-10-932-182A-1107 Sequence 1107, Ap
24	42	3.5	1884	18	US-11-079-463-2536 Sequence 2536, Ap
25	38	3.2	544	11	US-10-301-480-7940 Sequence 7940, Ap
26	38	3.2	544	12	US-10-301-480-7940 Sequence 7940, Ap
27	37.8	3.2	538	7	US-09-925-065A-342135 Sequence 342135, Appl
28	37.8	3.2	562	12	US-10-301-480-414923 Sequence 414923, Appl
29	37.8	3.2	562	12	US-10-301-480-414923 Sequence 414923, Appl
C 30	37.8	3.2	170995	17	US-11-121-086-35 Sequence 35, Appl
31	36.8	3.1	1624	10	US-10-131-826A-181 Sequence 181, App
32	36.8	3.1	1624	11	US-10-973-115B-181 Sequence 181, App
33	36.8	3.1	1624	13	US-10-137-873A-181 Sequence 181, App
34	36.8	3.1	1624	13	US-10-152-370-181 Sequence 181, App
35	36.8	3.1	1624	18	US-11-290-153-181 Sequence 181, App
C 36	36.6	3.1	487	7	US-09-925-065A-792220 Sequence 792220, Appl
C 37	36.6	3.1	487	7	US-09-925-065A-793983 Sequence 793983, Appl
C 38	36.6	3.1	487	7	US-09-925-065A-850079 Sequence 850079, Appl
39	36.6	3.1	538	7	US-09-925-065A-342136 Sequence 342136, Appl
40	36.6	3.1	538	7	US-09-925-065A-342137 Sequence 342137, Appl
41	36.6	3.1	562	12	US-10-301-480-414924 Sequence 414924, Appl
42	36.6	3.1	562	12	US-10-301-480-414925 Sequence 414925, Appl
43	36.6	3.1	562	12	US-10-301-480-1028333 Sequence 1028333, Appl
44	36.6	3.1	562	12	US-10-301-480-1028334 Sequence 1028334, Appl
C 45	36.6	3.1	1936	7	US-09-925-065A-83687 Sequence 83687, A

#### ALIGNMENTS

RESULT 1  
US-11-151-601-19  
; Sequence 19, Application US/11151601  
; Publication No. US200600003413A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel B.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Welch, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,  
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF  
; FILE REFERENCE: MF100-054PIRCPIOMNIDVIM  
; CURRENT APPLICATION NUMBER: US/11/151,601  
; CURRENT FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: US 10/170,789  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/882,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/212,078  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/934,406  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US01/26052  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,740  
; PRIOR FILING DATE: 2000-08-21

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(2874)
US-11-151-601-19

Query Match      86.9%; Score 1040; DB 17; Length 3143;
Best Local Similarity 89.1%; Pred. No. 1.1e-269;
Matches 1197; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCAGCTGTAGCTAGTCAAGTACGTAA 60
DB 1796 ATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCAGCTGTAGCTAGTCAAGTACGTAA 1855
QY 61 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCATTCTTGTGTCATCAGTC 120
DB 1856 ATCTGTGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCATTCTTGTGTCATCAGTC 1915
QY 121 AGCACTGTGACTCTTTTATAGTAAGTATAGTACCAAGAGTCCACACTGTGTGCC 180
DB 1916 AGCACTGTGACTCTTTTATAGTAAGTATAGTACCAAGAGTCCACACTGTGTGCC 1975
QY 181 TTTACAGCTATCAAGTCTGTGAGATGACCGAAGTTCGAAACAAAGAAATTTTGGGCCA 240
DB 1976 TTTACAGCTATCAAGTCTGTGAGATGACCGAAGTTCGAAACAAAGAAATTTTGGGCCA 2035
QY 241 CCATTTTGGGATTCAGCAGGTCTCTTCTGACTATACCTCTCCAGAAATTTCTCTTTTG 300
DB 2036 CCATTTTGGGATTCAGCAGGTCTCTTCTGACTATACCTCTCCAGAAATTTCTCTTTTG 2095
QY 301 AAGTACTACTGATTTACATTGTATGGGATCTCTACAGCCTCATGATCTACAGCCTG 360
DB 2096 AAGTACTACTGATTTACATTGTATGGGATCTCTACAGCCTCATGATCTACAGCCTG 2155
QY 361 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCA----- 403
DB 2156 GAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCAAGTGGTGGTAATA 2215
QY 404 ----- 403
DB 2216 ATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGGAATACCTAGCCTCTTAGTGTATGTGG 2275
QY 404 ----- 403
DB 2276 TTGTAGTGATAGACACAGGGATCCTGTCCAGGAGGCTTAAATTTGAAGGCGCCTTTA 2335
QY 404 -----GGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCCAAATATCTAGCTT 453
DB 2336 AATATAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCCAAATATCTAGCTT 2395
QY 454 CTGGATATGATTTTCATTTAGTATAGATCTGTGGGATCCACGGCTGTCTCTATGGAGGAT 513
DB 2396 CTGGATATGATTTTCATTTAGTATAGATCTGTGGGATCCACGGCTGTCTCTATGGAGGAT 2455
QY 514 ACTCTCCCTGTATGGGCAATTAATGACAGGTACAGATATCTTCAGGGTTGCTATTGCTGGGG 573
DB 2456 ACTCTCCCTGTATGGGCAATTAATGACAGGTACAGATATCTTCAGGGTTGCTATTGCTGGGG 2515
QY 574 CCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 633
DB 2516 CCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAACGTTATATGGGTCAAC 2575
QY 634 CTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATCGCAAGCAGAAAGTTCC 693
DB 2576 CTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATCGCAAGCAGAAAGTTCC 2635
QY 694 CCTCTGAACCAAACTGTTTACTGCTCTTACATGCTTCTTACATGCTTCTTGGATGAGAATGTCATTTTG 753
```

```
DB 2636 CCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTTGATGAGAAATGTCCATTTTG 2695
QY 754 CACATACAGTATATATCTGAGTTTATAGTGAGGCTGGAAAGCCATATGATTTACAGA 813
DB 2696 CACATACAGTATATATCTGAGTTTATAGTGAGGCTGGAAAGCCATATGATTTACAGA 2755
QY 814 TCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCGGGAGAAACATTTATGAACCTGC 873
DB 2756 TCTATCTCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCGGGAGAAACATTTATGAACCTGC 2815
QY 874 ATCTTTTGCACCTACCTTTCAAGAAAACCTTGGATCAGTATTTGCTCTCTTAAAGTGATAT 933
DB 2816 ATCTTTTGCACCTACCTTTCAAGAAAACCTTGGATCAGTATTTGCTCTCTTAAAGTGATAT 2875
QY 934 AATTTTGCACCTGTGTAGAACTCTCTGTATACACTGGCTATTTAACCAATGAGGAGGTT 993
DB 2876 AATTTTGCACCTGTGTAGAACTCTCTGTATACACTGGCTATTTAACCAATGAGGAGGTT 2935
QY 994 TAATCAACAGAAAACACAGAAATTTGATCATCATCATTTTGTATCTGCCATGTAAACATCTAC 1053
DB 2936 TAATCAACAGAAAACACAGAAATTTGATCATCATCATTTTGTATCTGCCATGTAAACATCTAC 2995
QY 1054 TCCTGAAAATTAATGTGGTGGCCATCGAGGGGTCTAGGTTTGTGGTAGTAACTTAATACC 1113
DB 2996 TCCTGAAAATTAATGTGGTGGCCATCGAGGGGTCTAGGTTTGTGGTAGTAACTTAATACC 3055
QY 1114 TTAACCCACATCTCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATA 1173
DB 3056 TTAACCCACATCTCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATA 3115
QY 1174 AGAATTAATAAAAAA----- 1197
DB 3116 AGAATTAATAAAAAA----- 3139

RESULT 2
US-11-151-601-21
; Sequence 21, Application US/11151601
; Publication No. US2006003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: WPI00-054P1RCP1QWNIQVIM
; CURRENT APPLICATION NUMBER: US/11/151,601
; PRIOR FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
```



Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 21  
; LENGTH: 2649  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-151-601-21

Query Match 65.0%; Score 778; DB 17; Length 2649;  
Best Local Similarity 86.4%; Pred. No. 4.7e-199;  
Matches 935; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 1 ATTTGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGTAGTACGTTACGTTAA 60  
Db ATTTGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGTAGTACGTTACGTTAA 1627

QY 61 ATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCCTGCTGCATCAGTC 120  
Db ATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCCTGCTGCATCAGTC 1687

QY 121 AGCAGTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCC 180  
Db AGCAGTGTGACTCTTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCC 1747

QY 181 TTACCAAGCTATCAAGTCTCAAGATGACCCCACTTGCAAAACAAAGGAATTTGGGCCA 240  
Db TTACCAAGCTATCAAGTCTCAAGATGACCCCACTTGCAAAACAAAGGAATTTGGGCCA 1807

QY 241 CCATTTTGGATTACAGAGGTCCTCTCTGACTATATCTCTCAGAAATTTCTCTTTTG 300  
Db CCATTTTGGATTACAGAGGTCCTCTCTGACTATATCTCTCAGAAATTTCTCTTTTG 1867

QY 301 AAGTACTACTGATTTACATGTATGGATGCTCTACAGCCTCATGCTACAGCCTG 360  
Db AAGTACTACTGATTTACATGTATGGATGCTCTACAGCCTCATGCTACAGCCTG 1927

QY 361 GAAGAAATTCCTACTGCTGCTTATATATATGCTGCTCA----- 403  
Db GAAGAAATTCCTACTGCTGCTTATATATATGCTGCTCA----- 1987

QY 404 ----- 403  
Db ----- 2047

QY 404 ----- 403  
Db ----- 2107

QY 404 -----GGTCAAAATAGAAATGACGATCAGGTGGAAGACTCCAAATATCTAGCTT 453  
Db AATATAAAATGGGTCAAAATAGAAATGACGATCAGGTGGAAGACTCCAAATATCTAGCTT 2167

QY 454 CTGATATGATTTTCAATGACTATAGATCGTGTGGGCATCCAGCCTGGTCTATGAGGAT 513  
Db CTGATATGATTTTCAATGACTATAGATCGTGTGGGCATCCAGCCTGGTCTATGAGGAT 2227

QY 514 ACCTCTCCCTGATGCTAATGACAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 573  
Db ACCTCTCCCTGATGCTAATGACAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGG 2287

QY 574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACCGGAACGTTATATGGGTACCC 633  
Db CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACCGGAACGTTATATGGGTACCC 2347

QY 634 CTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCC 693  
Db CTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCC 2407

QY 694 CCTCTGAACCAATCGTTTACTGTCTTTACATGTTTCTTGGATGAGATGTCATTGTTG 753  
Db CCTCTGAACCAATCGTTTACTGTCTTTACATGTTTCTTGGATGAGATGTCATTGTTG 2467

QY 754 CACATACAGTATATTAAGTGTGAGGGCTGAAAGCCATATGATTTTACAGA 813  
Db CACATACAGTATATTAAGTGTGAGGGCTGAAAGCCATATGATTTTACAGA 2527

QY 814 TCTATCCTCAGGAGACACAGCATATAGAGTTCTGTAATCGGAGAACATTATGAACCTGC 873  
Db TCTATCCTCAGGAGACACAGCATATAGAGTTCTGTAATCGGAGAACATTATGAACCTGC 2587

QY 874 ATCTTTTGCACCTCTTCAAGAAAACCTTGATGATCAGCTATTTGCTCTTAAAGTGTATAT 933  
Db ATCTTTTGCACCTCTTCAAGAAAACCTTGATGATCAGCTATTTGCTCTTAAAGTGTATAT 2647

QY 934 AA 935  
Db 2648 AA 2649

RESULT 3  
US-09-925-065A-669313/c  
; Sequence 669313, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 669313  
; LENGTH: 1346  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-669313

Query Match 9.9%; Score 118; DB 7; Length 1346;  
Best Local Similarity 100.0%; Pred. No. 3.2e-21;  
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 256 CAGGTCTCTTCTCCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAT 315  
Db 118 CAGGTCTCTTCTCCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAT 59

QY 316 TTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCC 373  
Db 58 TTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCC 1

RESULT 4  
US-09-925-065A-743558/c  
; Sequence 743558, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147

```
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 743558
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743558

Query Match
Best Local Similarity 9.5%; Score 113.6; DB 7; Length 609;
Matches 116; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 695 CTCGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAAATGTCATTTCG 754
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 205 CTCGAGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAAATGTCATTTCG 146
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 755 ACATACCAAGTATATTACTGAGTATTTTGTAGTGAGGGCTGGAAAGCCATATGATTTACAGAT 814
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 145 ACATACCAAGTATATTACTGAGTATTTTGTAGTGAGGGCTGGAAAGCCATATGATTTACAGT 86
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-09-925-065A-743559/c
; Sequence 743559, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 10827,135
; Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 743559
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743559

Query Match
Best Local Similarity 9.5%; Score 113.6; DB 7; Length 609;
Matches 116; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 695 CTCGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAAATGTCATTTCG 754
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 205 CTCGAGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAAATGTCATTTCG 146
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 755 ACATACCAAGTATATTACTGAGTATTTTGTAGTGAGGGCTGGAAAGCCATATGATTTACAGAT 814
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 145 ACATACCAAGTATATTACTGAGTATTTTGTAGTGAGGGCTGGAAAGCCATATGATTTACAGT 86
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-11-079-463-1186
; Sequence 1186, Application US/11079463
```

```
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRAC
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1186
; LENGTH: 2238
; TYPE: DNA
; ORGANISM: B.fragilis
US-11-079-463-1186

Query Match
Best Local Similarity 6.7%; Score 80.2; DB 18; Length 2238;
Matches 252; Conservative 0; Mismatches 233; Indels 12; Gaps 2;

Qy 402 CAGGGTCAAAATGAAATTCAGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATAT 461
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1747 CTGGGTGTAAGAAGCTAAAGACAGGTGGAAGCTGCCAANTATCTGG---GTGACTG 1803
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 462 GATTTTCATGACTTAGATCGTGTGGCCTCCACGGCTGGTCTTATGAGGATACCTCTCC 521
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1804 CTTATGTGGCAAGGACGTAATGGTAATCTGGGGATGGAGTTTCGGCGGATATATGACC 1863
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 522 CTGATGGCATTAATCAGAGGTGAGATATCTTCAGGGTGTCTATTCGTGGGGCCCCAGTC 581
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1864 ATCATGATGATGAGCGAAGGTACACCCGTGTTTAAAGCCGAGGTGCTGTGGCCGACCT 1923
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 582 ACTCTGTGGATCTTCTATGATACAGGATACACGAAACGTTTATATGGGTTCACCTGACCA 641
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1924 ACAGACTGGAATATTACGATACAGTATATACCGAACGCTTTATGCGCAGCGCGCAAGAA 1983
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 642 AATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGAA 701
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1984 AATGCCGAGGCTATTAAGACGACTTCAGCATTCAGCCGTGCAGAACCTGTCATGTTAAC 2043
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 702 CCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAAATGTCATTTGACATACC 761
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2044 C-----TGCTCTTGTACACGGTATGCGCAGATGATAATCTTCACATTCAGAACTGT 2094
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 762 AGTATATTACTGAGTATTTTGTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCT 821
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2095 ACAGAAATATGACAGACACCTGTTACAACTCGGAAACAGTTCGATATGAGGTATACACC 2154
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 822 CAGGAGACACAGCATAAGAGTTCTGGAATCGGAGAACATTATGAACTGCATCTTTTG 881
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2155 AACCGGAATCATGCACTCTATGTTGGAAATACCCGTAACCCACTTGTATACGAAGCTGACG 2214
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 882 CACTACCTTCAAGAAA 898
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 2215 AACTTCTCCGGAATAA 2231
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-09-925-065A-818446/c
; Sequence 818446, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 10827,135
; Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
```

;; PRIOR APPLICATION NUMBER: US 60/252,147  
;; PRIOR FILING DATE: 2000-11-20  
;; PRIOR APPLICATION NUMBER: US 60/250,092  
;; PRIOR FILING DATE: 2000-11-30  
;; PRIOR APPLICATION NUMBER: US 60/261,766  
;; PRIOR FILING DATE: 2001-01-16  
;; PRIOR APPLICATION NUMBER: US 60/289,846  
;; PRIOR FILING DATE: 2001-05-09  
;; NUMBER OF SEQ ID NOS: 957086  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 818446  
;; LENGTH: 612  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-925-065A-818446

Query Match 6.1%; Score 73.4; DB 7; Length 612;  
Best Local Similarity 98.7%; Pred. No. 2.5e-09;  
Matches 74; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 740 GAATGTCATTGTCACATACCAGTATATTAAGTATTTAGTGAGGGCTGGAAAGCC 799  
DB 612 GAATGTCATTGTCACATACCAGTATATTAAGTATTTAGTGAGGGCTGGAAAGCC 553  
QY 800 ATATGATTACAGAT 814  
DB 552 ATATGATTACAGGT 538

RESULT 8

US-11-208-288-1  
;; Sequence 1, Application US/11208288  
;; Publication No. US20060051366A1  
;; GENERAL INFORMATION:  
;; APPLICANT: CHANG, Chiwen  
;; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF  
;; FILE REFERENCE: 39533-0001  
;; CURRENT APPLICATION NUMBER: US/11/208,288  
;; PRIOR FILING DATE: 2005-08-18  
;; PRIOR APPLICATION NUMBER: US 60/605,013  
;; PRIOR FILING DATE: 2004-08-26  
;; NUMBER OF SEQ ID NOS: 8  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 1  
;; LENGTH: 3332  
;; TYPE: DNA  
;; ORGANISM: homo sapiens  
US-11-208-288-1

Query Match 5.8%; Score 70; DB 18; Length 3332;  
Best Local Similarity 47.7%; Pred. No. 3.9e-08;  
Matches 268; Conservative 0; Mismatches 290; Indels 4; Gaps 2;  
QY 402 CAGGTCATAATAGAAATTGAGATCAGGTGGGAAGGACTCCATATCTAGCTTCTCGATAT 461  
DB 1792 CTGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAAATG 1848  
QY 462 GATTTTCATTGACTTAGATCGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCC 521  
DB 1849 GGAATTTGGGACAAACAAACGAATTTGGGGCTGGTCTATATGGAGGGTACGTAACC 1908  
QY 522 CTGATGGCATTAATAGCAGAGTCTAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCCAGTC 581  
DB 1909 TCAATGGTCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGATAGCCGTGGCGCTGTA 1968  
QY 582 ACTCTGTGGATCTTTCTATGATACAGGATACACGGAAAGTGTATATGGGTACCCCTGACCCAG 641  
DB 1969 TCCCGGTGGGAGTACTATGACTCAGTGTACACAGAAGTGTATATGGGTCTCCCAACTCCA 2028  
QY 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAAGCAGAGAAAGTTCCTCCCTCTGAA 701  
DB 2029 GAAGCAACCTTGAACCATTAAGAAATTCACAGTCTATGAGCAGAGCTGAAATTTTAA 2088

RESULT 9

US-10-501-035-34  
;; Sequence 34, Application US/10501035  
;; Publication No. US20060046249A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Bristol-Myers Squibb Company  
;; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING  
;; FILE REFERENCE: D0185 PCT  
;; CURRENT APPLICATION NUMBER: US/10/501,035  
;; CURRENT FILING DATE: 2004-07-09  
;; PRIOR APPLICATION NUMBER: US 60/350,061  
;; PRIOR FILING DATE: 2002-01-18  
;; NUMBER OF SEQ ID NOS: 795  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 34  
;; LENGTH: 3407  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-501-035-34

Query Match 5.8%; Score 70; DB 11; Length 3407;  
Best Local Similarity 47.7%; Pred. No. 4e-08;  
Matches 268; Conservative 0; Mismatches 290; Indels 4; Gaps 2;  
QY 402 CAGGTCATAATAGAAATTGAGATCAGGTGGGAAGGACTCCATATCTAGCTTCTCGATAT 461  
DB 1867 CTGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAAATG 1923  
QY 462 GATTTTCATTGACTTAGATCGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCC 521  
DB 1924 GGAATTTGGGACAAACAAACGAATTTGGGGCTGGTCTATATGGAGGGTACGTAACC 1983  
QY 522 CTGATGGCATTAATAGCAGAGTCTAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCCAGTC 581  
DB 1984 TCAATGGTCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGATAGCCGTGGCGCTGTA 2043  
QY 582 ACTCTGTGGATCTTTCTATGATACAGGATACACGGAAAGTGTATATGGGTACCCCTGACCCAG 641  
DB 2044 TCCCGGTGGGAGTACTATGACTCAGTGTACACAGAAGTGTATATGGGTCTCCCAACTCCA 2103  
QY 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATCAAGCAGAGAAAGTTCCTCCCTCTGAA 701  
DB 2104 GAAGCAACCTTGAACCATTAAGAAATTCACAGTCTATGAGCAGAGCTGAAATTTTAA 2163  
QY 702 CCAATTCGTTTACTTGCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGCACATACC 761  
DB 2164 CAAAGTTGAGTACCTTCTCTTATTCATGGNAACAGCAGATGTAACGTTCACTTTCAGCAGTCA 2223  
QY 762 AGTATATTACTGAGTTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCT 821

Db 2224 GCTCAGATCTCCAAAGCCCTGTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCT 2283  
Qy 822 CAGGAGAGACACAGCATAAGAGTCTCTGAATCGGAGAACATTAATGAACCTGCACTTTTG 881  
Db 2284 GATGAAGACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2343  
Qy 882 CACTTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAAGTGATATAATTTGA 941  
Db 2344 CACTTCATAAAACAATGTTCTCTTTACCT-TAGCACCTCAAAATACCATGCCATTAA 2402  
Qy 942 CCTGTGTAGAACTCTCTGGTAT 963  
Db 2403 GCTTATTAAAACTCATTTTGT 2424

RESULT 10  
US-11-208-288-3  
; Sequence 3, Application US/11208288  
; Publication No. US20060031366A1  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, Chiwen  
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF  
; FILE REFERENCE: 39533-0001  
; CURRENT APPLICATION NUMBER: US/11/208,288  
; CURRENT FILING DATE: 2005-08-18  
; PRIOR APPLICATION NUMBER: US 60/605,013  
; PRIOR FILING DATE: 2004-08-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2217  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-208-288-3

Query Match 5.7%; Score 68.4; DB 18; Length 2217;  
Best Local Similarity 48.9%; Pred. No. 9.1e-08;  
Matches 214; Conservative 0; Mismatches 221; Indels 3; Gaps 1;  
Qy 402 CAGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAT 461  
Db 1708 CTGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA--CAATTTTCAAAAATG 1764  
Qy 462 GATTTCATTCAGCTAGATCGTGTGGCATCCAGCGCTGCTCTATGGAGGATACCTCTCC 521  
Db 1765 GGATTTGTGACAAACAACGAATTGCAATTTGGGGCTGGTCATATGGAGGTACGTAAAC 1824  
Qy 522 CTGATGGCATTAATGCGAGGTCAGATATCTTCAGGGTGTCTATGCTGGGGCCCAAGTC 581  
Db 1825 TCAATGGTCTGGGATCGGAGTGAAGTGGCGTGTCAAGTGTGGNATAGCGTGGCGCTGTA 1884  
Qy 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCAACCTGACGAG 641  
Db 1885 TCCCGGTGGGAGTACTATGACTCAGGTGATACAGAACGTTTACATGGGTCCTCCAACTCCA 1944  
Qy 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAA 701  
Db 1945 GAAGACAACCTTGACCAATTAAGAAAATTCAAAGTCATGAGCAGAGCTGAAAATTTTAA 2004  
Qy 702 CCAATCGTTTACTGCTCTTACATGTTTCTGATGAAATGTCATTTTGGCAATACC 761  
Db 2005 CAAGTTAGTACCTCTCTTATTCATGGAACAGCAGATGATAAGCTTCACTTTTCCAGCAGTCA 2064  
Qy 762 AGTATATTACTAGTGTGTGTGTGAGGGTGGAAAGCCATATGATTTTACAGATCTATCCT 821  
Db 2065 GCTCAGATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTACT 2124  
Qy 822 CAGGAGAGACACAGCATA 839  
Db 2125 GATGAAGACCATGGAATA 2142

RESULT 11  
US-10-522-789-1  
; Sequence 1, Application US/10522789  
; Publication No. US20050260732A1  
; GENERAL INFORMATION:  
; APPLICANT: TANABE SEIYAKU CO., LTD.  
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV  
; FILE REFERENCE: 03-039-PCT  
; CURRENT APPLICATION NUMBER: US/10/522,789  
; CURRENT FILING DATE: 2005-01-28  
; PRIOR APPLICATION NUMBER: US 60/398,761  
; PRIOR FILING DATE: 2002-07-29  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 2301  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2301)  
; OTHER INFORMATION:  
US-10-522-789-1

Query Match 5.7%; Score 68.4; DB 10; Length 2301;  
Best Local Similarity 48.9%; Pred. No. 9.2e-08;  
Matches 214; Conservative 0; Mismatches 221; Indels 3; Gaps 1;  
Qy 402 CAGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAT 461  
Db 1792 CTGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA--CAATTTTCAAAAATG 1848  
Qy 462 GATTTCATTCAGCTAGATCGTGTGGCATCCAGCGCTGCTCTATGGAGGATACCTCTCC 521  
Db 1849 GGATTTGTGACAAACAACGAATTTGGGCTGGTCATATGGAGGTACGTAAAC 1908  
Qy 522 CTGATGGCATTAATGCGAGGTCAGATATCTTCAGGGTGTCTATGCTGGGGCCCAAGTC 581  
Db 1909 TCAATGGTCTCGGATCGGGAAGTGGCGTGTTCAGAGTGTGGAATAGCGTGGCGCTGTA 1968  
Qy 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCAACCTGACGAG 641  
Db 1969 TCCCGGTGGGAGTACTATGACTCAGGTGATACAGAACGTTTACATGGGTCTCCCAACTCCA 2028  
Qy 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAA 701  
Db 2029 GAAGACAACCTTGACCAATTAAGAAAATTCAAAGTCATGAGCAGAGCTGAAAATTTTAA 2088  
Qy 702 CCAATCGTTTACTGCTCTTACATGTTTCTGATGGAATGTCATTTTGGCAATACC 761  
Db 2089 CAAGTTGAGTACCTCTCTTATTCATGGAACAGCAGATGATAAGCTTCACTTTTCCAGCAGTCA 2148  
Qy 762 AGTATATTACTAGTGTGTGTGTGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCCT 821  
Db 2149 GCTCAGATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTACT 2208  
Qy 822 CAGGAGAGACACAGCATA 839  
Db 2209 GATGAAGACCATGGAATA 2226

RESULT 12  
US-10-505-928-476  
; Sequence 476, Application US/10505928  
; Publication No. US2006008532A1  
; GENERAL INFORMATION:  
; APPLICANT: Ludwig Institute for Cancer Research et al.  
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
; FILE REFERENCE: 28967/39178  
; CURRENT APPLICATION NUMBER: US/10/505,928  
; CURRENT FILING DATE: 2004-08-27  
; PRIOR APPLICATION NUMBER: US 60/363,019

```
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 476
; LENGTH: 2788
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-476

Query Match      5.4%; Score 64.4; DB 9; Length 2788;
Best Local Similarity 48.1%; Pred. No. 1.2e-06;
Matches 182; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 462 GATTTCATTGACTTAGATTCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCC 521
Db 2039 GGTTTCATTGATGAAAAAAGATAGCCATATGGGGCTGGTCTCTATGGAGGATACGTTTCA 2098

QY 522 CTGATGCAATTAATGCAGAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTC 581
Db 2099 TCACCTGGCCCTTGCATCTGGAACCTGCTCTTTTCAATGTTGGTATAGCAGTGGCTCCAGTC 2158

QY 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAAACGTTATATGGGTACACCTCGACAG 641
Db 2159 TCCAGCTGGGAATATTACGGCTCTGTCTACACAGAGAGATTTCATGGGTCTCCCAACAAG 2218

QY 642 AATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAA 701
Db 2219 GATGATAATCTTGAGCACTATAAGAATTCAACTGTGATGGCAAGAGCAGAATATTTTCA 2278

QY 702 CCAATCGTTTACTGCTCTTACATGGTTCCTGGATGAGAATGTCCATTTTGCACATACC 761
Db 2279 AATGTAGACTATCTTCTCATCCAGCAACAGCAGATGATAATGTGCACCTTTCAAAACTCA 2338

QY 762 AGTATATTACTAGATTTTTTTAGTGGGGCTGGAAAGCCATATGATTTACAGATCTATCCT 821
Db 2339 GCACAGATTGCTAAAGCTCTGTTAATGCAAGTGGATTTCCAGGCAATGTGGTACTCT 2398

QY 822 CAGGAGACACACAGCATA 839
Db 2399 GACCAGAACCCAGGCTTA 2416

RESULT 13
US-11-186-284-54
; Sequence 54, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP001-029P22NM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 2814

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Homo sapiens fibroblast activation protein, alpha
; OTHER INFORMATION: (FAP), mRNA
US-11-245-147-168

Query Match      5.4%; Score 64.4; DB 17; Length 2814;
Best Local Similarity 48.1%; Pred. No. 1.2e-06;
Matches 182; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 462 GATTTCATTGACTTAGATTCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCC 521
Db 2039 GGTTTCATTGATGAAAAAAGATAGCCATATGGGGCTGGTCTCTATGGAGGATACGTTTCA 2098

QY 522 CTGATGCAATTAATGCAGAGGTTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTC 581
Db 2099 TCACCTGGCCCTTGCATCTGGAACCTGCTCTTTTCAATGTTGGTATAGCAGTGGCTCCAGTC 2158

QY 582 ACTCTGTGGATCTTCTATGATACAGGATACAGGAAACGTTATATGGGTACACCTCGACAG 641
Db 2159 TCCAGCTGGGAATATTACGGCTCTGTCTACACAGAGAGATTTCATGGGTCTCCCAACAAG 2218

QY 642 AATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAA 701
Db 2219 GATGATAATCTTGAGCACTATAAGAATTCAACTGTGATGGCAAGAGCAGAATATTTTCA 2278

QY 702 CCAATCGTTTACTGCTCTTACATGGTTCCTGGATGAGAATGTCCATTTTGCACATACC 761
Db 2279 AATGTAGACTATCTTCTCATCCAGCAACAGCAGATGATAATGTGCACCTTTCAAAACTCA 2338

QY 762 AGTATATTACTAGATTTTTTTAGTGGGGCTGGAAAGCCATATGATTTACAGATCTATCCT 821
Db 2339 GCACAGATTGCTAAAGCTCTGTTAATGCAAGTGGATTTCCAGGCAATGTGGTACTCT 2398

QY 822 CAGGAGACACACAGCATA 839
Db 2399 GACCAGAACCCAGGCTTA 2416

RESULT 14
US-11-245-147-168
; Sequence 168, Application US/11245147
; Publication No. US20060030541A1
; GENERAL INFORMATION:
; APPLICANT: GARCIA, TERESA
; APPLICANT: ROMAN ROMAN, SERGIO
; APPLICANT: BARON, ROLAND
; APPLICANT: CALL, KATHERINE
; APPLICANT: THEILHABER, JOACHIM
; APPLICANT: CONNOLLY, TIMOTHY
; APPLICANT: JACKSON, AMANDA
; APPLICANT: BUSHNELL, STEVEN
; APPLICANT: RAWADI, GEORGES
; TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE
; FILE REFERENCE: 37991-0023
; CURRENT APPLICATION NUMBER: US/11/245,147
; CURRENT FILING DATE: 2005-10-07
; PRIOR APPLICATION NUMBER: PCT/IB02/02211
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/281,400
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 168
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Homo sapiens fibroblast activation protein, alpha
; OTHER INFORMATION: (FAP), mRNA
US-11-245-147-168
```

```
Query Match      5.4%; Score 64.4; DB 18; Length 2814;
Best Local Similarity 48.1%; Pred. No. 1.2e-06;
Matches 182; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 462 GATTTCATTGACCTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCC 521
Db 2039 GGTTCATTGATGAAAAAGAAATAGCCATATGGGCTGGTCTCTATGGAGGATACGTTTCA 2098

QY 522 CTGATGGCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGTGGGGCCCCAGTC 581
Db 2099 TCACCTGGCCCTTGCACTGGAACCTGCTCTTTTCAAATGGGTATAGCAGTGGCTCCAGTC 2158

QY 582 ACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACCAAG 641
Db 2159 TCCAGCTGGGAATATTACGGCTCTGTCTACAGAGAGATTTCATGGGTCTCCCAACAAAG 2218

QY 642 AATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGAA 701
Db 2219 GATGATAATCTTGAGCACTATAAGAAATTCAACTGTGATGGCAAGAGCAGAATATTTCCAG 2278

QY 702 CCAATCGTTTACTGCTCTTACATGGTTTCTCTGGATGAGAAATGTCATTTTGCACATACC 761
Db 2279 AATGTAGACTATCTTCTCATCCAGGAACAGCAGATGATATGTGCACCTTCAAAACTCA 2338

QY 762 AGTATATTACTGAGTTTATTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTT 821
Db 2339 GCACAGATTGCTAAGCTCTGGTTAATGCAAGTGGATTTCAGGCNAATGTGGTACTCT 2398

QY 822 CAGGAGAGACACAGCATTA 839
Db 2399 GACCAGAAACCACGGCTTA 2416

RESULT 15
US-11-226-869-428
; Sequence 428, Application US/11226869
; Publication No. US2006069054A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C14
; CURRENT APPLICATION NUMBER: US/11/226,869
; CURRENT FILING DATE: 2005-09-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-226-869-428

Query Match      5.2%; Score 61.8; DB 18; Length 535;
Best Local Similarity 58.4%; Pred. No. 3.2e-06;
Matches 108; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 462 GATTTCATTGACCTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCC 521
Db 273 GGTTCATTGATGAAAAAGAAATAGCCATATGGGCTGGTCTCTATGGAGGATACGTTTCA 332

QY 522 CTGATGGCATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGTGGGGCCCCAGTC 581
Db 333 TCACCTGGCCCTTGCACTGGAACCTGCTCTTTTCAAATGGGTATAGCAGTGGCTCCAGTC 392

QY 582 ACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACCAAG 641
Db 393 TCCAGCTGGGAATATTACGGCTCTGTCTACAGAGAGATTTCATGGGTCTCCCAACAAAG 452

QY 642 AATGA 646
Db 453 GATGA 457
```

Search completed: May 4, 2006, 09:26:07  
Job time : 1637 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	502	23.5	931	2	T32919		hypothetical prote
2	484.5	22.6	738	2	A87516		dipeptidyl peptida
3	398.5	18.6	741	2	JCS142		X-Pro dipeptidyl-p
4	379	17.7	711	2	S66261		X-Pro dipeptidyl-p
5	359.5	16.8	793	2	T41703		dipeptidyl aminope
6	342.5	16.0	760	1	S23752		dipeptidyl-peptida
7	340.5	15.9	792	1	A39914		dipeptidyl-peptida
8	337.5	15.8	766	1	CDHU26		dipeptidyl-peptida
9	314	14.7	931	2	A49737		dipeptidyl aminope
10	310	14.5	818	1	A30107		dipeptidyl aminope
11	303.5	14.2	779	2	T25173		hypothetical prote
12	303.5	14.2	799	2	T25174		hypothetical prote
13	284.5	13.3	829	2	T19514		hypothetical prote
14	262	12.2	803	2	A41793		dipeptidyl aminope



```
QY 123 CACTGTGACTTCTTTATAGTAAGTATAGTATACCAAGAG-----AATCCACAC 170
Db 594 ---Cys-----PheYsSerGluArgAlaAenGlyLysLeuAlaLeuAapLeuAapHis 610
QY 171 TGTGTGTCTTACAGCTATCA-----ATCTCTGAAGATGACCCCACTTGCAAAACA 224
Db 611 GlyPheAlaCysTyrMetThrSerValGlySerProAlaGlu-----CysArgPhe 627
QY 225 AAGGAATTT---TGG-----GCCACCAT 245
Db 628 TyrSerPheArgTrpLysGluAenGluValLeuProSerThrValTyrAlaAlaAenIle 647
QY 246 TTGGATTACAGAGGTCTCT---CTTCTCTGAC-----TATACTCTCTCCAGAAATTTTCTCT 296
Db 648 ThrValSerGlyHisProGlyGlnProAapLeuHisPheAapSerProGluMetIleGlu 667
QY 297 TTGAAAGT---ACTACTGTGATTACATTGTATGGATGCTCTACAGCCTCATGACTA 353
Db 668 PheGlnSerLysLysThrGlyLeuMetHisTyrAlaMetIleLeuAupProSerAenPhe 687
QY 354 CAGCTGGAAGAAATATCTACTGTGCTGTTCATATATGTGTCTCT----- 401
Db 688 AspProTyrLysLysTyrProValPheHisTyrValTyrGlyGlyProGlyIleGlnIle 707
QY 401 ----- 401
Db 708 ValHisAenAapPheSerTrpIleGlnTyrIleArgPheCysArgLeuGlyTyrValVal 727
QY 401 ----- 401
Db 728 ValPheIleAapAenArgGlySerAlaHisArgGlyIleGluPheGluArgHisIleHis 747
QY 402 -----CAGGTCAAATAGAAATAGCATGAGTGTGGAAGCACTCCAATATCTAGCTTCT 455
Db 748 LysLysMetGlyThrValGluValGluAapGlnValGluGlyLeuGlnMetLeuAlaGlu 767
QY 456 CGATAT---GATTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGA 512
Db 768 ArgThrGlyGlyPheMetAapMetSerArgValValValHisGlyTrpSerTyrGlyGly 787
QY 513 TACTCTCTCCGTGATCAATTAATGACAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGG 572
Db 788 TyrMetAlaLeuGlnMetIleAlaLysHisProAenIleTyrArgAlaIleAlaGly 807
QY 573 GCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCAC 632
Db 808 GlyAlaValSerAapTrpArgLeuTyrAapThrAlaTyrThrGluAargTyrMetGlyTyr 827
QY 633 CCTGACCAGAAATCAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTC 692
Db 828 Pro---LeuGluGluHisValTyrGlyAlaSerSerIleThrGlyLeuValGluLysLeu 846
QY 693 CCCTCTGMAACCAATCGTTTACTGCTCTTACATGTTCTCTGGATGAGATGTCCTATTTT 752
Db 847 ProAapGluProAenAargMetLeuValHisGlyLeuMetAapGluAenValHisPhe 866
QY 753 GCATACACAGTATATATCTAGTATTTTATGAGGGCTGGAAGCCATATGATTTACAG 812
Db 867 AlaHisLeuThrHisLeuValAapGluCysIleLysLysGlyLysTrpHisGluLeuVal 886
QY 813 ATCTATCTCAGAGAGACACAGCATAAAGATGTTCTCGAATCGGGAGAACATATGAACTG 872
Db 887 IlePheProAenGluArgHisGlyValArgAenAenAapAlaSerIleTyrLeuAapAla 906
QY 873 CATCTTTTGACCTTCCAAAGAA 896
Db 907 ArgMetMetTyrPheAlaGlnGln 914
```

## RESULT 2

A87516

dipeptidyl peptidase IV [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: A87516  
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.J.;  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: A87516  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-738 <SFO>  
A;Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI00000C7616; GB:AE005673; NID:g13423647; P;  
C;Genetics:  
A;Gene: CC2154

## Alignment Scores:

Pred. No.:	2,22e-37	Length:	738
Score:	484.50	Matches:	106
Percent Similarity:	48.3%	Conservative:	53
Best Local Similarity:	32.2%	Mismatches:	109
Query Match:	22.6%	Indels:	61
DB:	2	Gaps:	6

US-10-825-632-4 (1-1197) x A87516 (1-738)

```
QY 3 TTTCAGGCGCACAAAGACTCCCTTTAGAGCATCATCTGTAGTACTGAGTTCAGTTACGTAAAT 62
Db 394 PheSerAlaSerIleAapThrProIleGluArgArgLeuTyrGluValSerTyrAlaLys 413
QY 63 CTGTGAGAGGTGCACAGGCTGACTGACCTGGCTACTCACATTTCTTGTGTCGATCAGTCAG 122
Db 414 ProGlyLysProLysAlaLeuThrSerAlaGlyGlyTyrTrpAlaAlaLysValAlaAap 433
QY 123 CACTGTGACTTCTTTTATAAGTATAGTATAACAGGAAGAATCCACACTGTGTGTCCTT 182
Db 434 AsnGlyAlaPheAlaGlyThrTyrSerAapProLysThrProSerGlnThrAlaLeu 453
QY 183 TACAAGGTATCAAGTCTCT-----GAAGATGACCCCACTTGCAAAACAAAG 227
Db 454 TyrSerAlaAapGlyLysArgValArgTrpIleGluGluAenLysLeuAlaGluGlyHis 473
QY 228 GAATTTTGGGCCACCATTTTGGATTGAGCAGGTCTCTTCTCTGACTATATCTCTCCAGAA 287
Db 474 ProTyrTrp-----ProTyrAlaAlaAenLeuProGln-----ProGlu 486
QY 288 ATTTTCTCTTTTGAAGTACTACTGTGATTACATTGTATGGATGCTCTACAAGCCTCAT 347
Db 487 PheGlySerLeuLysAlaAlaAapGlyGluThrLeuHisTyrGluIleLeuLysProIle 506
QY 348 GATCTACAGCCTGGAAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCAG--- 404
Db 507 GlyPheAapProAlaLysLysTyrProAlaIleValSerValTyrGlyGlyProHisAla 526
QY 404 ----- 404
Db 527 GlnArgValMetLysAenTrpHisSerProSerGluArgThrTyrLeuGluAlaGlyTyr 546
QY 404 ----- 404
Db 547 ValIlePheLysLeuAapAenArgGlySerGlyAenArgSerAlaLysPheMetArgAla 566
QY 405 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAATATCTA 449
Db 567 LeuAapArgLysLeuGlyThrValGluValGluAapGlnLeuLeuLysPhePheLeu 586
QY 450 GCTTCTCGATATGATTTCAATTGACTTAGATCGTGTGGCATCCACGGCTGCTCTATGGA 509
Db 587 AlaSerGln---ProTyrValAapAlaAapLysLeuGlyValMetGlyTrpSerTyrGly 605
QY 510 GGATACCTCTCCCTGATGGCATTAATGCAGAGGTGCAGATATCTTCAGGGTTGCTATTGCT 569
Db 606 GlyPheMetAlaLeuMetLeuLeuThrAlaGluAenThrProPheLysAlaGlyAlaAla 625
```

```
QY 570 GGGGCCCCAGTCACTCTGTGGAGCTTCTATGATACAGGATACACGGAACTGTTATATGGT 629
|||||
Db 626 GlyAlaProProThrGluTrpSerLeuThrAlaValThrGluArgTyrMetGly 645
|||||
QY 630 CACCTGACCAAGTGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAG 689
|||||
Db 646 LysProAspGluAenLysAlaGlyTyrAlaTyrSerAspAlaAenAsnArgIleAspLys 665
|||||
QY 690 TTCCCTCTGAACCAAAATCGTTTACTCTCTATACATGGTTTCTCGATGAGAAATGCCAT 749
|||||
Db 666 Leu-----AlaProGlySerLeuLeuLeuLeuHisGlyMetAlaAspAsnValIle 683
|||||
QY 750 TTTCACATACACGATATATACGATGTTTATAGTACGGCTGGAAAGCCATATATTA 809
|||||
Db 684 PheGluAenSerThrArgLeuMetAlaAlaLeuGlnArgLysAlaIleLeuPheGluMet 703
|||||
QY 810 CAGATCTATCTCAGGAGACACAGC 836
|||||
Db 704 AlaMetTyrProGlyGluArgHisSer 712
|||||
RESULT 3
JC5142
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Xanthomonas maltophilia
C:Species: Xanthomonas maltophilia
C>Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5142
R:Kabashima, T.; Ito, K.; Yoshimoto, T.
J. Biochem. 120, 1111-1117, 1996
A:Title: Dipeptidyl peptidase IV from Xanthomonas maltophilia: Sequencing and expression
A:Reference number: JC5142; MUID:97164011; PMID:9010758
A:Accession: JC5142
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <KAB>
A:Cross-references: UNIPROT:P95782; UNIPARC:UPI0000085F8A; DDBJ:D83263; NID:g1753196; PI
C:Comment: This enzyme catalyzes the hydrolytic removal of amino-terminal dipeptidyl res
C:Superfamily: dipeptidyl-peptidase IV
F:4-18/Domain: transmembrane #status predicted <TM>
F:610/Active site: Ser #status predicted
F:685/Active site: Asp #status predicted
F:717/Active site: His #status predicted
Alignment Scores:
Pred. No.: 3.06e-29 Length: 741
Score: 398.50 Matches: 98
Percent Similarity: 44.3% Conservative: 57
Best Local Similarity: 28.0% Mismatches: 120
Query Match: 18.6% Indels: 75
DB: 2 Gaps: 8
US-10-825-632-4 (1-1197) x JC5142 (1-741)
QY 3 TTGAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGTACGTAAAT 62
|||||
Db 396 PheArgAlaGlyIleGluSerAlaArgGluSerGlnIleTyrAlaValPro--LeuGln 414
|||||
QY 63 CTGGAGAGGTGACAGCTGACGTGACCGTGGCTACTCATCTCTGTCGATCAGTACG 122
|||||
Db 415 GlyGlyGlnProGlnArgLeuSerLysAlaProGlyMetHisSerAlaSerPheAlaArg 434
|||||
QY 123 CACTGTGACTCTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTT 182
|||||
Db 435 AenAlaSerValTyrValAspSerTrpSerAsnAsnSerThrProGlnIleGluLeu 454
|||||
QY 183 TACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCCACC 242
|||||
Db 455 PheArgAlaAenGly-----GluLysIleAlaThr 464
|||||
QY 243 ATTTTG-----GATTCAAGAGTCTCTCTCTCGACTATACT- 278
|||||
Db 465 LeuValGluAenAspLeuAlaAspProLysHisProTyrAlaArgTyrArgGluAlaGln 484
|||||
```

```
QY 279 ---CCTCAGCAAAATTTTCTCTTTGAAAGTACTACTCGATTTTACA---TTGTATGGGATG 332
|||||
Db 485 ArgProValGluPheGlyThrLeuThrAlaAlaAspGlyLysThrProLeuAenTyrSer 504
|||||
QY 333 CTCCTACAGCTCATGATCTACAGCTCGGAAAGAAATATCTCTACTGTGTGTTTCATATAT 392
|||||
Db 505 ValIleLysProAlaGlyPheAspProAlaLysArgTyrProValAlaValTyrValTyr 524
|||||
QY 393 GGTGTGCTCT----- 401
|||||
Db 525 GlyGlyProAlaSerGlnThrValThrAspSerTrpProGlyArgGlyAspHisLeuPhe 544
|||||
QY 401 ----- 401
Db 545 AenGlnTyrLeuAlaGlnGlnGlyTyrValValPheSerLeuAspAsnArgGlyThrPro 564
|||||
QY 402 -----CAGGTCAAATAGAAATTGAC 422
|||||
Db 565 ArgArgGlyArgAspPheGlyAlaLeuTyrGlyLysGlnGlyThrValGluValAla 584
|||||
QY 423 GATCAGGTGGAAGGACTCCAATATCTAGCTTCGATATGATTTCAATTCAGCTTAGCT 482
|||||
Db 585 AspGlnLeuArgGlyValAlaIleTrpLeu---LysGlnGlnProTrpValAspProAlaArg 603
|||||
QY 483 GTGGCAATCCAGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGG 542
|||||
Db 604 IleGlyValGlnGlyTyrSerAsnGlyGlyTyrMetThrLeuMetLeuLeuAlaLysAla 623
|||||
QY 543 TCAGATATCTTCAGGGTGTCTATGCTGGGCCCCAGTCACTCTGTGGATCTCTCTATGAT 602
|||||
Db 624 SerAspSerTyrAlaCysGlyValAlaGlyValProValThrAspTrpGlyLeuTyrAsp 643
|||||
QY 603 ACAGGATACACGGAACGTTATATGGGTCACCTGACAGAAATGACCAACAGGCTTACTTATA 662
|||||
Db 644 SerHisTyrThrGluArgTyrMetAspLeuProAlaArgAsnAspAlaGlyTyrArgGlu 663
|||||
QY 663 GGATCTGTGGCCATGCAAGCAGCAAAAGTTCCTCTGTAACCAAAATCGTTTACTGCTTTA 722
|||||
Db 664 AlaArgValLeuThrHisIleGluGlyLeuArgSerPro-----LeuLeuLeuLeu 680
|||||
QY 723 CATGTTTCTCGGATGAGATGTCATTTCACATACACCATATATTACTGAGTGTTTTA 782
|||||
Db 681 HisGlyMetAlaAspAspAsnValLeuPheThrAsnSerThrSerLeuMetSerAlaLeu 700
|||||
QY 783 GTGAGGCTGCAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAAGA 842
|||||
Db 701 GlnLysArgGlyGlnProPheGluLeuMetThrTyrProGlyAlaLysHisGlyLeuSer 720
|||||
QY 843 GTTCTCTGAATCGGAGAACATTTATGAACCTG 872
|||||
Db 721 GlyAlaAspAlaLeuHisArgTyrArgVal 730
|||||
RESULT 4
S66261
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
C:Species: Flavobacterium meningosepticum
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S66261
R:Kabashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A:Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV
A:Reference number: S66261; MUID:95314307; PMID:7793970
A:Accession: S66261
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-711 <KAB>
A:Cross-references: UNIPROT:Q47900; UNIPARC:UPI0000082115; EMBL:D42121; NID:g577283; PID:
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase
Alignment Scores:
Pred. No.: 2.12e-27 Length: 711
```



```
QY 465 -----TTCAATTGACTTAGTACGTCGTGGGCATCCACGGCTGGTCTCTATGGA 509
Db 630 TrpAlaaspLeuProPheValaspGluAsnHisValGlyIleTrpGlyTrpSerTyrgly 649
QY 510 GGATACCTCTCCCTGAGGATTAAGTACAGAGGTACAGATATCTTCAGGGTTCATTGCT 569
Db 650 GlyTyrglyLeuThrLeuLysThrLeu---GluThrGlnAspValPheSerTyrglyMetAla 668
QY 570 GGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTATATGGGT 629
Db 669 ValAlaProValThrAspTrpArgLeuTyrglyAspSerValTyrglyGluArgTyrglyMetAsp 688
QY 630 CACCTGACAGGATGAACAGGCTATTACTTAGGATCTGTGGCCATCCAGCAGAGAAAG 689
Db 689 LeuProGlnTyrglyAsnLysGluGly---TyrLysAsnSerGlnIleHisAspTyrglyLys 707
QY 690 TTCCCTCTGAAACCAATCCCTTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCAT 749
Db 708 Phe---LysGlnLeuLysArgPhePheValAlaHisGlyThrGlyAspAsnValHis 726
QY 750 TTTCACATACCATATATTACTAGAGTATTTTATAGTGGGCTGGA---AAGCCATATCAT 806
Db 727 PheGlnHisSerMetHisLeuMetAspGlyLeuAsnLeuAlaAsnCysTyrglyAsnTyrgly 746
QY 807 TTACAGATCTATCTCAGGAGACACAGCATAAAGTTCCTGGAATCGGAGAACATATAT 866
Db 747 MetAlaValPheProAspSerAlaHisSerIle-----SerTyrglyHisAsnAla 762
QY 867 GAATGATCTTTTGGACTACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTA 923
Db 763 SerLeuSerIleTyrglyHisArgLeuSerGluTrpIleGlyAlaLeuGlyArgIle 781

RESULT 6
S23752
dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
N:Alternate names: CD26 alpha subunit; THAM alpha subunit
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S23752; A46465; A56030
R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
J. Biol. Chem. 267, 2200-2208, 1992
A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di
A:Reference number: S23752; MUID:92129288; PMID:1370813
A:Accession: S23752
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-760 <MAR>
A:Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384
R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;
J. Immunol. 147, 447-454, 1991
A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase
A:Reference number: A46465; MUID:91302787; PMID:1712807
A:Accession: A46465
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <VIV>
A:Cross-references: UNIPARC:UPI0000172A2E
A:Experimental source: M14.T thymoma cells, Swiss nu/nu
A>Note: sequence extracted from NCBI backbone (NCBI:42236)
R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
Biochemistry 33, 15204-15214, 1994
A:Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.
A:Reference number: A56030; MUID:95032780; PMID:7999781
A:Accession: A56030
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 746-760 <BER>
A:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620
C:Genetics:
A:Gene: CD26
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
F:213,223,315,514,679/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
F:624,702,734/Active site: Ser, Asp, His #status predicted

Alignment Scores: 6.13e-24 Length: 760
Pred. No.: 342.50 Matches: 101
Score: 42.1% Conservative: 51
Best Local Similarity: 28.0% Mismatches: 120
Query Match: 16.0% Indels: 89
DB: 1 Gaps: 14

US-10-825-632-4 (1-1197) x S23752 (1-760)

QY 15 AAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGTTACGTTAAATCCTCGAGAGGTG 74
Db 417 LysGluMetProGlyGlyArgAsnLeuTyrglyLysIle----- 428
QY 75 ACAAGGCTGACTGACCGTGGCTACTCACATTCTTGTGTCATCAGT----- 119
Db 429 ---GlnLeuThrAsp-----HisThrAsnValLysCysLeuSerCysAspLeuAsnPro 445
QY 120 CAGCAGCTGCACTTCTTTATAGTAAGTATAGTAACAGAGAATCCACAC----- 170
Db 446 GluArgCysGlnTyrglyAlaValSerPheSerLysGluAlaLysTyrglyGlnLeuGly 465
QY 171 TGT-----GTTGTCCTTTTCAAGCTATCATCAAGTCTCGAAGATGACCCAACTTGC 218
Db 466 CysTrpGlyProGlyLeuProLeuTyrglyThrLeuHisArgSerThrAspHisLysGluLeu 485
QY 219 AAAACAAGGAATTTTGGGCCACCATTTGGATTTCAGCAGGT-----CCTCTTCT 269
Db 486 ArgValLeuGlu-----AspAsnSerAlaLeuAspArgMetLeuGln 499
QY 270 GACTATATCTCTCCAGAA-----ATTTTCTCTTTTGAAGTACTACTGGATTACA 320
Db 500 AspValGlnMetProSerLysLysLeuAspPheIleValLeuAsnGluThrArgPhe--- 518
QY 321 TTGTATGGGATGCTCTACAAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTG 380
Db 519 TrpTyrglyMetIleLeuProProHis---PheAspLysSerLysTyrglyProLeuLeu 537
QY 381 CTGTTCATATATGGTGGTCTCT----- 401
Db 538 LeuAspValTyrglyAlaGlyProCysSerGlnLysAlaAspAlaSerPheArgLeuAsnTrp 557
QY 401 ----- 401
Db 558 AlaThrTyrglyLeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySer 577
QY 402 -----CAGGGTCAATAGAAATT 419
Db 578 GlyTyrglyGlnGlyAspLysIleMetHisAlaIleAsnArgArgLeuGlyThrLeuGluVal 597
QY 420 GACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGAT 479
Db 598 GluAspGlnIleGluAlaAlaArgGlnPheVal---LysMetGlyPheValAspSerLys 616
QY 480 CGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGCTATTATGCAG 539
Db 617 ArgValAlaIleTrpGlyTrpSerTyrglyGlyTyrglyValThrSerMetValLeuGlySer 636
QY 540 AGGTGAGATATCTTCAGGGTTCATATGCTGGGGCCCCAGTCACTCTGTGGATCTTCTAT 599
Db 637 GlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGluTyrgly 656
QY 600 GATACAGATACACGGAACCTTATATGGGT-----CACCTGTACCAAGATGAACAGGGC 653
Db 657 AspSerValTyrglyThrGluArgTyrglyMetGlyLeuProIleProGluAspAsnLeuAspHis 676
QY 654 TATTACTTAGGATCTGTGGCCATGCGAAGCAGAAAGTTCCTCTCGAACCAATCGTTTA 713
Db 677 TyrArgAsnSerThrValMetSerArgAlaGluHisPhe-----LysGlnValGluTyrgly 694
QY 714 CTGCTCTTACATGGTTCCTGGATGAGAATGTCCATTTTGGACATACCATGATATTACTG 773
```

```

Db 695 LeuLeuIleHisGlyThrAlaAspAspAenValHisPheGlnGlnSerAlaGlnIleSer 714
Qy 774 AGTTTTAGTGGGCGGAAGCCATGATGATTTACAGATCTATCTCTCAGGAGACAC 833
Db 715 LysValLeuValAspAlaGlyValAspPheGlnAlaMetTrpTyrThrAspGluAspHis 734
Qy 834 AGCATAAGAGTCTCTGAATCGGAGAACATTTATGAACATCTTTTGCACTACTCTCAA 893
Db 735 GlyIleAlaSerSerThrAlaHisGlnHisIleTyrSerHisMetSerHisPheLeuGln 754
Qy 894 GAA 896
Db 755 Gln 755

RESULT 7
A39914
N:Alternate names: GP110; membrane glycoprotein 110K; OX-61
N:Contains: dipeptidyl-peptidase IV, soluble form
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C:Accession: A39914; A33315; B33315; A60730; A42203; S38949; A31781
R:Hong, W.; Doyle, D.
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987
A:Title: cDNA cloning for a bile canalicular domain-specific membrane glycoprotein of rat
A:Reference number: A39914; MUID:88068516; PMID:3479775
A:Accession: A39914
A:Molecule type: mRNA
A:Residues: 1-792 <HON>
A:Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN:
R:Ogata, S.; Misumi, Y.; Ikehara, Y.
J. Biol. Chem. 264, 3596-3601, 1989
A:Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA an
A:Reference number: A33315; MUID:89123496; PMID:2563382
A:Accession: A33315
A:Molecule type: mRNA
A:Residues: 1-37, 'A', 39-182, 'I', 184-331, 'T', 333-351, 'C', 353-393, 'V', 395-561, 'L', 563-623,
A:Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:G203973; PIDN:AAA41096.1; PID:
A:Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332
A:Accession: B33315
A:Molecule type: protein
A:Residues: 1-20:35-54; 427-443:505-509; 511-520:530-538; 593-600:602-608; 618-627 <HO2>
A:Cross-references: UNIPARC:UPI0000172A37
A36: UNIPARC:UPI0000172A37
R:McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.
Hepatology 11, 534-544, 1990
A:Title: Identification of the bile canalicular cell surface molecule GP110 as the ectop
quence.
A:Reference number: A60730; MUID:90228896; PMID:1970322
A:Accession: A60730
A:Molecule type: protein
A:Residues: 28-47, 'XX', 50-53, 55-58 <MCC>
A:Cross-references: UNIPARC:UPI0000172A38
R:Ogata, S.; Misumi, Y.; Teuji, E.; Takami, N.; Oda, K.; Ikehara, Y.
Biochemistry 31, 2582-2587, 1992
A:Title: Identification of the active site residues in dipeptidyl peptidase IV by affini
A:Reference number: A42203; MUID:92190188; PMID:1347701
A:Accession: A42203
A:Molecule type: protein
A:Residues: 'R', 625-630, 'X', 632-648 <OG2>
A:Cross-references: UNIPARC:UPI0000172A39
R:Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993
A:Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl p
A:Reference number: S38949; MUID:94128239; PMID:7905271
A:Accession: S38949
A:Status: preliminary
A:Molecule type: protein
A:Residues: 281-302 <IWA>
A:Cross-references: UNIPARC:UPI0000172A3A
R:Hong, W.; Doyle, D.
J. Biol. Chem. 263, 16892-16898, 1988

```

```

A:Title: Membrane orientation of rat gp110 as studied by in vitro translation.
A:Reference number: A31781; MUID:89034185; PMID:3182821
A:Accession: A31781
A:Molecule type: mRNA
A:Residues: 1-40 <HO3>
A:Cross-references: UNIPARC:UPI0000172A3B
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidyl-peptidase, membrane-bound form #status experimental <MATM>
F:1-792/Product: dipeptidyl-peptidase, membrane-bound form #status experimental <MATM>
F:1-28/Domain: signal sequence #link MATS #status experimental <SIG>
F:1-6/Domain: intracellular #status predicted <INT>
F:1-28/Domain: transmembrane #status predicted <TMN>
F:29-792/Domain: extracellular #status predicted <EXT>
F:35-792/Product: dipeptidyl-peptidase, soluble form #status experimental <PRO>
F:83-90/148-217/227-319-521-686/Binding site: carboxylate (Asn) (covalent) #status predi
F:631/Active site: Ser #status experimental
F:709,741/Active site: Asp, His #status predicted
Alignment Scores:
Pred. No.: 9,56e-24 Length: 792
Score: 340.50 Matches: 103
Percent Similarity: 42.9% Conservative: 51
Best Local Similarity: 28.7% Mismatches: 120
Query Match: 15.9% Indels: 85
DB: 1 Gaps: 15
US-10-825-632-4 (1-1197) x A39914 (1-792)
Qy 15 AAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTAAATCTCGGAGAGTG 74
Db 424 LysGluMetProGlyGlyArgAsnLeuTyrLysIle 435
Qy 75 ACAAGGCTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGT 119
Db 436 ---GlnLeuThrAsp-----HisThrAsnLysLysCysLeuSerCysAspLeuAsnPro 452
Qy 120 CAGCAGCTGTCACATCTCTT-----ATAAGTAAGTAGTAGTACAGCAAGAACATCCACAC 170
Db 453 GluArgCysGlnTyrTyrSerValSerLeuSerLysGluAlaLysTyrTyrGlnLeuGly 472
Qy 171 TGT-----GTGTCCTTTTACAAGCTATCAAGTCTCGAGAGTCTCTCTCGATATACT 218
Db 473 CysArgGlyProGlyLeuProLeuTyrThrLeuHisArgSerThrAspGlnLysGluLeu 492
Qy 219 AAAACAAAGAAATTTGGGCCCCATTTTGGATTTCAGCAGGTCTCTCTCGATATACT 278
Db 493 ArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAspValGln 509
Qy 279 CTCCAGAA-----ATTTCTCTTTGAAGTACTACTGGATTTACATTTGATGGG 329
Db 510 MetProSerLysLysLeuAspPheIleValLeuAsnGluThrArgPhe---TrpTyrGln 528
Qy 330 ATGCTCTACAGCTCATGATCTACAGCTGGAAAGAAATATCTCTGCTGCTCATTA 389
Db 529 MetIleLeuProHis---PheAspLysSerLysLysTyrProLeuLeuIleAspVal 547
Qy 390 TATGGTGGTCTCT----- 401
Db 548 TyrAlaGlyProCysSerGlnLysAlaAspAlaAlaPheArgPheAsnTrpAlaThrTyr 567
Qy 401 ----- 401
Db 568 LeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGln 587
Qy 402 -----CAGGGTCAAAATAGAAATGACGATCAG 428
Db 588 GlyAspLysIleMethHisAlaIleAsnLysArgLeuLeuGluValGluAspGln 607
Qy 429 GTGGAAGGAGCTC---CAATATCTTAGCTTCTCGATATGATTTTCATTTAGATCGGTG 485
Db 608 IleGluAlaAlaArgGlnPheLeu-----LysMetGlyPheValAspSerLysGlnVal 625

```





Db 551 CysSerGlnLysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThr 570  
QY 401 -----  
Db 571 GluAsnIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIle 590  
QY 402 -----CAGGTCAAATAGAAATGACGATCAGGTGGAAGGACTC 440  
Db 591 MeHisAlaIleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAla 610  
QY 441 CAATATCTAGCTTCTCGATATGATTTCAATGACTTGTAGTCGTGGGCGATCCACGGCTGG 500  
Db 611 ArgGlnPhe---SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrp 629  
QY 501 TCCTATGGAGATACCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGGTT 560  
Db 630 SerTyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCys 649  
QY 561 GCTATTCTGGGCGCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGT 620  
Db 650 GlyIleAlaValAlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArg 669  
QY 621 TATATGGT-----CACCTGACCAAGATGAACAGGCTATTACTAGGATCTGTGGCC 674  
Db 670 TyrMetGlyLeuProThrProGluAspAsnLysTyrArgAsnSerThrValMet 689  
QY 675 ATCGACGACAAAGTTCCCTCTGAACCAATCGTTTACTCTTACATGTTTCCCTG 734  
Db 690 SerArgAlaGluAsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAla 707  
QY 735 GATGAGATCTCCATTTGCGACATGACAGTATATTCAGTATTTTTCAGGGCTGGA 794  
Db 708 AspAspAsnValHisPheGlnSerAlaGlnIleSerLysAlaLeuValAspValGly 727  
QY 795 AAGCATATGATTACAGATCTATCTCTCAGGAGACACACAGCATAGAAGTCTCGAATCG 854  
Db 728 ValAspPheGlnAlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAla 747  
QY 855 GGAAACCATATTGAATGATCTTTTGCACTACCTTCAAGAA 896  
Db 748 HisGlnHisIleTyrThrHisMetSerHisPheIleLysGln 761  
RESULT 9  
A49737  
dipeptidyl aminopeptidase (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein O5045; protein YOR219c; protein YOR50-9  
C:Species: Saccharomyces cerevisiae  
C:Date: 03-Mar-1994 #sequence revision 03-Mar-1994 #text\_change 09-Jul-2004  
C:Accession: A49737; S45451; S60946; S67112; S71721  
R:Santa Anna-A, S.; Herskowitz, I.  
submitted to the Protein Sequence Database, July 1993  
A:Reference number: A49737  
A:Accession: A49737  
A:Molecule type: DNA  
A:Residues: 1-931 <SAN>  
A:Cross-references: UNIPROT:P33894; UNIPARC:UPI0000136060; GB:L21944; NID:g347196; PIDN:  
R:Anna-Arriola, S.S.; Herskowitz, I.  
Yeast 10, 801-810, 1994  
A:Title: Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase  
A:Reference number: S45451; MUID:95066382; PMID:7975897  
A:Accession: S45451  
A:Molecule type: DNA  
A:Residues: 1-931 <ANN>  
A:Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:g347196; PIDN:AAA35119.1; PI  
R:Gallison, F.; Dujon, B.  
submitted to the EMBL Data Library, October 1995  
A:Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome  
A:Reference number: S60938  
A:Accession: S60946  
A:Molecule type: DNA  
A:Residues: 1-931 <GAL>  
A:Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; F  
R:Boyer, J.; Fairhead, C.; Gallion, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon, B.

submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67104  
A:Accession: S67112  
A:Molecule type: DNA  
A:Residues: 1-931 <BOY>  
A:Cross-references: UNIPARC:UPI0000136060; EMBL:275127; NID:g1420507; PIDN:CAA99437.1; PI  
A:Experimental source: strain S288C  
R:Gallison, F.; Dujon, B.  
Yeast 12, 877-885, 1996  
A:Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of  
A:Reference number: S71713; MUID:96437977; PMID:8840505  
A:Accession: S71721  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-931 <GAW>  
A:Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; PI  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
C:Genetics:  
A:Gene: SGD:STE13; YC11  
A:Cross-references: SGD:S0005745; MIPS:YOR219c  
A:Map position: 15R  
C:Function:  
A:Description: involved in processing of alpha-factor prepropheromone  
C:Keywords: dipeptidylpeptidase hydrolase; glycoprotein; transmembrane protein  
F:113-150/Domain: transmembrane #status predicted <TM>  
F:377/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:785,863,896/Active site: Ser, Asp, His #status predicted  
Alignment Scores:  
Pred. No.: 3,18e-21 Length: 931  
Score: 314.00 Matches: 100  
Percent Similarity: 39.4% Conservative: 54  
Best Local Similarity: 25.6% Mismatches: 133  
Query Match: 14.7% Indels: 104  
DB: 12 Gaps: 12  
US-10-825-632-4 (1-1197) x A49737 (1-931)  
QY 3 TTGAAGGACCAAGACATCCCTTTAGAGCATACCTGTAGTACGTAGTACGTACGTA--- 59  
Db 562 PheThrAlaAsnGluIleGlyValMetSerGlnHisLeuTyrSerIleSerLeuThrAsp 581  
QY 60 -----AACTCTGGAGAGGTGACAAAGCTGACT 86  
Db 582 SerThrThrGlnAsnThrPheGlnSerLeuGlnAsnPro-----Ser 595  
QY 87 GACCGTGGCTACTCACATCTCTGCTGCATCAGTACAGCAGTCTGCTTATTAAGTAAG 146  
Db 596 AsplysTyrAspPheTyrAspPheGluLeuSerSerSerAlaArgTyrAlaIleSerLys 615  
QY 147 TATAGTAACCAAGAAATCCA----- 167  
Db 616 LysLeuGlyProAspThrProIleLysValAlaGlyProLeuThrArgValLeuAsnVal 635  
QY 168 -----CACTGTGTGCTCCCTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTGC 218  
Db 636 AlaGluIleHisAspAspSerIleLeuGlnLeuThrLysAspGlu----- 650  
QY 219 AAAACAAGCAATTTTGGGCCACCATTTTGGATTCCAGCAGGTCTCTTCTCTGACTATACT 278  
Db 651 LysPheLysGlu-----LysIleLysAsnTyrAsp 660  
QY 279 CCTCCAGAAATTTCTTTTGAAGTACTACT-----GGATTTTACATTTGATGGG 329  
Db 661 LeuPro---IleThrSerTyrLysThrMetValLeuAspAspGlyValGluIleAsnTyr 679  
QY 330 ATGCTCTACAGGCTCATCATCTACAGCTCGAAGCAATATCTCTGCTGCTGTTCTATA 389  
Db 680 IleGluIleLysProAlaAsnLeuAsnProLysLysLysTyrPheIleLeuValAsnIle 699  
QY 390 TATGCTGCTCTCAGGGTCAA----- 410  
Db 700 TyrGlyGlyProGlySerGlnThrPheThrLysSerSerLeuAlaPheGluGlnAla 719



Qy 411 -----ATAGAAATTGACGATCAGGTGGAGGA----- 437  
Db 720 valValSerGlyLeuAspValIleValLeuGlnIleGluProArgGlyThrGlyGlyLys 739  
Qy 437 ----- 437  
Db 740 GlyTrpSerPheArgSerTrpAlaArgGluLysLeuGlyTyrrTrpGluProArgAspIle 759  
Qy 438 -----CTCCAATPATCTAGCTTCTCGATATGATTTTCATTGACTTAGATTCGTG 485  
Db 760 ThrGluValThrIlystysPheIleGlnArgAsnSerGlnHisIleAspGluSerLysIle 779  
Qy 486 GGCATCCAGCGCTGCTTATGAGGATACCTCTCCCTGATGGCATTA---ATGCAGAGG 542  
Db 780 AlaIleTrpGlyTrpSerTyrrGlyGlyPheThrSerLeuLysThrValGluLeuAspAsn 799  
Qy 543 TCAGATATCTCAGGTTCTTATCTGGGGCCCGAGTCACCTCTGTGATCTTCTATGAT 602  
Db 800 GlyAspThrPheLysTyrrAlaMetAlaValAlaProValThrAsnTrpThrLeuTyrrAsp 819  
Qy 603 ACAGGATACACGGAACGTTATATGGTCCACCTGACAGAAATGAACAGGGCTATTACTTA 662  
Db 820 SerValTyrrThrGluArgTyrrMetAsnGlnProSerGluAsnHisGluGlyTyrrPheGlu 839  
Qy 663 GGAATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTA 722  
Db 840 ValSerThrIleGlnAsnPhelysserPheGluSer---LeuLysArgLeuPheIleVal 858  
Qy 723 CATGGTTCTCGATGAGAAATGCCATTTTGCATACCACTATATTACTGAGTTTCTTA 782  
Db 859 HisGlyThrPheAspAspAsnValHisIleGlnAsnThrPheArgLeuValAspGlnLeu 878  
Qy 783 GTGAGCGCTGGA---AAGCCATATGATTACAGATCTATCTCTCAGGAGACACAGATA 839  
Db 879 AsnLeuLeuGlyLeuThrAsnTyrrAspMetHisIlePheProAspSerAspHisserIle 898  
Qy 840 AGAGTTCCTGAATCGGGAGAACATTATGACTGCATCTTTTGCATCTACCTACCTCAAGAAAC 899  
Db 899 ArgTyrrHisAsnAlaGlnArgIleValPheGlnLysLeuTyrrTyrrLeuArgAspAla 918  
Qy 900 CTGGATCATCGTATTCTGCTCTAAAAGTGATA 932  
Db 919 PheAlaGluArgPheAspAsnThrGluValLeu 929  
RESULT 10  
A30107  
diPeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein YHR028c  
C;Species: Saccharomyces cerevisiae  
C;Date: 07-Jun-1990 #sequence\_revision 30-May-1997 #text\_change 09-Jul-2004  
C;Accession: S46780; A30107  
R;Du, Z.  
submitted to the EMBL Data Library, June 1994  
A;Description: The sequence of *S. cerevisiae* cosmid 8082.  
A;Reference number: S46773  
A;Accession: S46780  
A;Molecule type: DNA  
A;Residues: 1-818 'DUZ'  
A;Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031A5F; EMBL:U10399; NID:g500689; PID  
R;Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.  
J. Cell Biol. 108, 1363-1373, 1999  
A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an in  
A;Reference number: A30107; MUID:89174971; PMID:2647766  
A;Accession: A30107  
A;Molecule type: DNA  
A;Residues: 1-82 'H', 84-124, 'N', 126-181, 'LRRLET', 189-199, 'N', 201-365, 'DFKRGKERKE', 376-57  
A;Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484  
A;Note: the authors translated the codon ACC for residue 572 as Asn  
C;Genetics:  
A;Gene: SGD:DAP2; STE13; MIPS:YHR028c  
A;Cross-references: SGD:S0001070; MIPS:YHR028c  
A;Map position: 8R

C;Superfamily: dipeptidyl-peptidase IV  
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vacu  
F;30-45/Domain: transmembrane #status predicted <TM>  
F;63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Alignment Scores: 7.41e-21 Length: 818  
Pred. No.: 310.00 Matches: 98  
Score: 37.4% Conservative: 38  
Best Local Similarity: 26.9% Mismatches: 122  
Query Match: 14.5% Indels: 106  
DB: 1 Gaps: 13

US-10-825-632-4 (1-1197) x A30107 (1-818)

Qy 3 TTTGAAGGCCAACAGACTCCCTTTAGACATCACTGTAGTAGTACGTAGTACGTAAAT 62  
Db 451 PheIleSerThrArgLysSerThrThrGluArgHisValTyrrTyrrIleAspLeuArgSer 470  
Qy 63 CTGGAGAGGTGACAGGCTGACTGACCGTGCTACTCA-----CATTCTTGCTGATC 116  
Db 471 ProAsnGluIleIleGluValThrAspThrSerGluAspGlyValTyrrAspValSerPhe 490  
Qy 117 ACTCAGCAGCTGCTACTCTTTATAGTAAGTATAGTAACACAGAAATCCACAC----- 170  
Db 491 SerSerGlyArgArgPheGlyLeuLeuThrTyrrLysGlyProLysValProTyrrGlnLys 510  
Qy 171 -----TGT-----GTG 176  
Db 511 IleValAspPheHisSerArgLysAlaGluLysCysAspLysGlyAsnValLeuGlyLys 530  
Qy 177 TCCCTTTTACAAAGTATCAAGTCTCTGAAGATGACCCCACTGTCAAACAAAGAAATTTTGG 236  
Db 531 SerLeuTyrrHisLeuGluLysAsnGlu----- 539  
Qy 237 GCCACCATTTTGGATTACAGAGTCTCTCTCTGCTACTACTCTCTCCAGAAATTTTCTCT 296  
Db 540 -----ValLeuThrLysIle-----LeuGluAspTyrrAlaValProArgLysSerPhe 555  
Qy 297 TTTGAAGTACTACTCGATTACATTGATGGG-----ATGCTCTACAAGCCTCATCAT 350  
Db 556 ArgGluLeuAsnLeuGlyLysAspGluPheGlyLysAspIleLeuValAsnSerTyrrGlu 575  
Qy 351 CTACAGCCT-----GGAAAGAAATATCTCTACTGCTGCTGTTCATA 389  
Db 576 IleLeuProAsnAspPheAspGluThrLeuSerAspHisTyrrProValPhePhePheala 595  
Qy 390 TATGGTGGTCTCTCAG-----GGTCAATAAGAAATTTGACAGATCAG 428  
Db 596 TyrGlyGlyProAsnSerGlnGlnValValLysThrPheSerValGlyPheAsnGluVal 615  
Qy 404 ----- 404  
Db 616 ValAlaSerGlnLeuAsnAlaIleValValValAspGlyArgGlyThrGlyPheLys 635  
Qy 405 -----GGTCAATAAGAAATTTGACAGATCAG 428  
Db 636 GlyGlnAspPheArgSerLeuValArgAspArgLeuGlyAspTyrrGluAlaArgAspGln 655  
Qy 429 GTGAAGGATCCCAATATCTAGTCTTCTCGATATGATGATTTTCATTGACTTAGATCGTGGGC 488  
Db 656 IleSerAlaAlaSerLeuTyrrGlySer--LeuThrPheValAspProGlnLysIleSer 674  
Qy 489 ATCCACCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATATCCAGAGGTCAGAT 548  
Db 675 LeuPheGlyTrpSerTyrrGlyGlyTyrrLeuThrLysThrLeuGluLysAspGlyGly 694  
Qy 549 ---ATCTTCAGGGTTCGATTGCTGGGGCCCGCCAGTCACCTCTGTGATCTTCTATGATACA 605  
Db 695 ArgHisPheLysTyrrGlyMetSerValAlaProValThrAspTrpArgPheTyrrAspSer 714  
Qy 606 GGATACACGGAACGTTATATGGTTCACCTGACCAAGAAATGAACAGGCTATTACTTAGGA 665

Db 715 ValTyrThrGluArgTyrMetHisThrProGlnGluAsnPheAspGlyTyrValGluSer 734  
Qy 666 TCCTGTG-----GCCATGCCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTA 713  
Db 735 SerValHisAsnValThrAlaLeu-----AlaGlnAlaAsnArgPhe 748  
Qy 714 CTCTCTTACATGGTTCTCGATGAGATGCCATTTTGGACATACCATGATATATCTG 773  
Db 749 LeuLeuMetHisGlyThrGlyAspAsnValHisPheGlnAsnSerLeuLeuPheLeu 768  
Qy 774 AGTTTTTTAGTGGGCTGGA--AAGCCATATGATTTACAGATCTATCTCCAGGAGAGA 830  
Db 769 AspLeuLeuAspLeuAsnGlyValGluAsnTyrAspValHisValPheProAspSerAsp 788  
Qy 831 CACAGCATAAGA 842  
Db 789 HisSerIleArg 792

RESULT 11  
T25173  
hypothetical protein T23F1.7a - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T25173  
R:Wilkinson, J.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19990  
A:Accession: T25173  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-779 <WIL>  
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN:CAB03411.1;  
A:Experimental source: clone T23F1  
C:Genetics: CSRP:T23F1.7a  
A:Map position: 5  
A:Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1  
C:Superfamily: dipeptidyl-peptidase IV

Alignment Scores:  
Pred. No.: 3,02e-20 Length: 779  
Score: 303.50 Matches: 93  
Percent Similarity: 42.8% Conservative: 62  
Best Local Similarity: 25.7% Mismatches: 118  
Query Match: 14.2% Indels: 89  
DB: 2 Gaps: 14

US-10-825-632-4 (1-1197) x T25173 (1-779)  
Qy 21 TCCCTTTAGAGCATCACCTGTACGTACGTACGTACGTAAATCTCGAGAGGTGACAAGG 80  
Db 431 AlaProLysProSerHisArgSerLeuPheSerTyr-----Ser 443  
Qy 81 CTGACTGACCGTGGCTACTCATCTTGTGCTGCATCAGT-----CAGCACTGTGAC 131  
Db 444 LeuAlaAspGluSerArgAsnSerAlaTyrCysIleSerCysSerIleLysAsnCysThr 463  
Qy 132 TTTCTTTATAGTATAGTATAGTAAACAGAGATCCACACTGTGTGCTTTCCTTACAGCTA 191  
Db 464 TrpAlaGlnAlaGlnMetAspAspGlnMetLysThrAlaIleValSerCysLysGlyPro 483  
Qy 192 TCAGTCTCTGAAGT-----GACCCCACTTGCACAAACAAAG 227  
Db 484 AlaAlaProHisThrAlaIleValAsnLeuThrArgMetAspSerAspLysLysThrGlu 503  
Qy 228 -----GAATTTTGGGCCACCATTTTGGATTCAGAGTCTCTCT 266  
Db 504 HisAlaAsnLeuLeuTyrAspLysThrTyrGlnAsnArgValGluGluAlaGly---Leu 522  
Qy 267 CTGACTATACCTCTCCAGAAATTTCTCTTTTGGAAAGTACTACTGGATTTCATGTTAT 326  
Db 523 ProValIleIleLysGluThrIle-----LysIleSerAspAspPheAlaLeu 539

Qy 327 GGGATGCTCTCAAGCCTCATGATCTACAGCTCGAAGAAGATAT-----CCTACT 377  
Db 540 IleLysLysSerIleProLysAspIleTyrAsnArgAspLysHisGlnAlaIleProLeu 559  
Qy 378 GTGCTGTTCATATATGGTGTCTCTCAG----- 404  
Db 560 IleValHisValTyrGlyGlyProAsnAspGlnAsnThrLysGluAlaThrGlnIleGly 579  
Qy 404 ----- 404  
Db 580 IleGluGluValValAlaSerAlaSerGlnAlaAlaIleLeuArgIleAspGlyArgGly 599  
Qy 405 -----GGCAATAGAA 416  
Db 600 SerGlyGlyArgGlyTrpLysTyrArgSerAlaIleTyrGlyGlnLeuGlyThrValGlu 619  
Qy 417 ATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAT---GATTTCATTGAC 473  
Db 620 ValGluAspGlnIleLysAlaIleLysValValLeuArgLeuTyrArgHisLeuLeuAsp 639  
Qy 474 TTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGCATTA 533  
Db 640 AlaArgArgValAlaValPheGlyTrpSerTyrGlyGlyPheMetThrLeuSerMetVal 659  
Qy 534 ATGCAGAGGTCAGAT---ATCTTCAGGTTGCTATTGCTGGGGCCCGCAGTCTCTGG 590  
Db 660 AsnGluAlaProGlnPhePheLysCysAlaValSerValAlaProValThrAsnPhe 679  
Qy 591 ATCTTCTATGATACAGATACACGGAACGTTATATGGTCCACCTGACCAAGATGAACAG 650  
Db 680 AlatyTyrAspAlaThrTyrThrGluArgTyrMetGly-----AspAlaProLeuGlu 697  
Qy 651 GGCTATTACTTAGATGCTGTGCCATGCAAGCAAAAGTTCCCTCTCGAACCAAAATCGT 710  
Db 698 SerTyr-----SerAspValThrLysLysLeuAspAsnPheLysSer-----ThrArg 713  
Qy 711 TTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGCACATACAGTATATTA 770  
Db 714 LeuLeuLeuMetHisGlyLeuLeuAspAspAsnValHisPheGlnAsnSerAlaIleLeu 733  
Qy 771 CTGAGTTTTTTAGTGGGCTCGAAGCCATATGATTTTACAGATCTATCTCTCAGAGAGA 830  
Db 734 IleAspGluLeuGlnAsnArgGlyValAspPheAspLeuMetValTyrProAsnGlnAla 753  
Qy 831 CACAGCATAGAGATTCTCTGAATCGGGAGAACATTATGAATGCACTCTTTTGCACCTT 890  
Db 754 HisSerLeuSerSerArgThrSer-----HisValValGlyLysMetThrHisPheLeu 771  
Qy 891 CAAGAA 896  
Db 772 ArgGln 773

RESULT 12  
T25174  
hypothetical protein T23F1.7b - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T25174  
R:Wilkinson, J.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19990  
A:Accession: T25174  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-799 <WIL>  
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN:CAB03412.1;  
A:Experimental source: clone T23F1  
C:Genetics: CSRP:T23F1.7b  
A:Map position: 5  
A:Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1  
C:Superfamily: dipeptidyl-peptidase IV

\_\_\_\_\_

Qy 404 -----

QY ----- 404 -----

----- 404

Db	633	ILsAlaTyRAlaArgIleAspValArgGlyThrGlyGlyArgGlyTyrAspValValysGlu	652
Qy	405	-----CGTCAATAGAAATGAGATCAGGTGGAGGACTCCCAATAT	446
Db	653	AlaValTyRArgIleLeuGlyAspAlaGluValValAspThrLeuAspMetIleArgAla	672
Qy	447	CTAGCTCTCGATATGATTTCATGACTTAGTCGTGGGCGCCACCGCTGGTCTAT	506
Db	673	PheIleAsnThrPheGlyPheIleAspGluAspArgIleAlaValMetGlyTyrPsrTyr	692
Qy	507	GGAGGTACCTC---TCCCTGATGCAATTAATGCAAGGTGAGATCTCTCAGGGTTGCT	563
Db	693	GlyGlyPheLeuThrSerLysIleAlaIleLysAspGlnGlyGluLeuValLysCysAla	712
Qy	564	ATTGCTGGGCGCCAGCTACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTAT	623
Db	713	IleSerIleAlaProValThrAspPheLysTyrTyrAspSerAlaTyrThrGluArgTyr	732
Qy	624	ATGGGTCACTCAGCAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCA	683
Db	733	LeuGlyGlnProAlaGluAsnLeuGlnGlyTyrIleAsnThrAsnValIleProHisAla	752
Qy	684	GAAGAATCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGAT	743
Db	753	ArgAsnValThrAsnVal-----LysTyrLeuLeuAlaHisGlyGluArgAspAsn	770
Qy	744	GTCCATTTTGCAATACACAGTATATTAAGTGTCTTACATGGTTTCTCGGATGAGAT	803
Db	771	ValHisTyrGlnAsnSerAlaArgTrpSerGluAlaLeuGlnGlnAsnGlyIleHisPhe	790
Qy	804	GATTACAGATCTATCTCAGAGAGACACACAGCATAGAGTCTCTGAATCGGAGAACAT	863
Db	791	ThrGlnLeuValTyrAlaAsnGluAlaHisSerLeu-----SerHis	804
Qy	864	TATGAACCTGATCTTTTGCACTACCTTCAAGAAACCTT	902
Db	805	LysLeuPheHisLeuTyrGlyGluValGlnArgPheLeu	817
RESULT 14			
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine			
C:Species: Bos primigenius taurus (cattle)			
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004			
C:Accession: A41793			
R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthold, R.J.; Shimasaki, S.			
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992			
A:Title: Differential expression of two distinct forms of mRNA encoding members of a dip			
A:Reference number: A41793; MUID:92108018; PMID:1729689			
A:Accession: A41793			
A:Status: preliminary; not compared with conceptual translation			
A:Molecule type: mRNA			
A:Residues: 1-803 <WAD>			
A:Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A83C; GB:M76429; NID:94087119; PIDN:			
A:Note: sequence extracted from NCBI Backbone (NCBI:75138)			
C:Superfamily: dipeptidyl-peptidase IV			
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein			
F:257,342/Binding site: carbohydrate (Asn) (covalent) #status predicted			
Alignment Scores:			
Pred. No.:	2,58e-16	Length:	803
Score:	262.00	Matches:	89
Percent Similarity:	37.6%	Conservative:	45
Best Local Similarity:	25.0%	Mismatches:	152
Query Match:	12.2%	Indels:	70
DB:	2	Gaps:	9
US-10-825-632-4 (1-1197) x A41793 (1-803)			
Qy	3	TTTGAAGGCAACCAAGATCCCTTTAGAGCATCATCCTGTAGCTAGTACGTAGTAAAT	62
Db	440	PheLeuSerThrGluAspLeuProArgArgGlnLeuTyrSerAlaSerThrVal---	458
RESULT 15			
I68600			
dipeptidyl aminopeptidase like protein - human			
C:Species: Homo sapiens (man)			
C:Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000			



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocorelation Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: May 2, 2006, 00:42:23 ; Search time 55.1366 Seconds  
(without alignments)  
3063.366 Million cell updates/sec

Title: US-10-825-632-4

Perfect score: 2140

Sequence: 1 attttgaaggcaccacaaagac.....ttactaaaaaaaaaaaaaa 1197

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/abs/ABSWEB\_spool/US10825632/runat\_01052006\_105946\_3226/app.query.fasta\_1  
-DBS=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p  
-USR=US10825632 @CGN 1 1 694 @runat\_01052006\_105946\_3226 -NCPU=6 -ICPU=3  
-NO\_MMAP -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DSV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBEXT=7

Database : UniProt 05.80:\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1645.5	76.9	898	1 DPP8 HUMAN	Q6v1x1 homo sapien
2	1605.5	75.0	892	1 DPP8_MOUSE	Q80ya7 mus musculus
3	1094.5	51.1	862	1 DPP9_MOUSE	Q8bvg4 mus musculus
4	1093.5	51.1	863	1 DPP9_HUMAN	Q8et12 homo sapien
5	1084.5	50.7	923	2 Q4SBM6_TETNG	Q4sbm6 tetraodon n
6	1059	49.5	847	2 Q6GR22_XENLA	Q6gr22 xenopus lae
7	1032.5	48.2	508	2 Q75273_HUMAN	Q75273 homo sapien
8	730.5	34.1	621	2 Q7PTT8_ANOGA	Q7ptt8 anopheles g
9	730.5	34.1	886	2 Q7QBK1_ANOGA	Q7qbki anopheles g
10	710.5	33.2	740	2 Q5VTK8_ANOGA	Q5vct8 anopheles g
11	704.5	32.9	1053	2 Q9VC20_DROME	Q9vc20 drosophila
12	704.5	32.9	1113	2 Q9VC19_DROME	Q9vc19 drosophila
13	639	29.9	557	2 Q5TXJ2_ANOGA	Q5txj2 anopheles g
14	599	28.0	803	2 Q54U01_DICDI	Q54u01 dictyosteli
15	502	23.5	927	2 Q965K3_CAEEL	Q965k3 caenorhabdi
16	502	23.5	931	2 Q44987_CAEEL	Q44987 caenorhabdi

17	484.5	22.6	738	2	Q9A6E0_CAUCR	Q9a6e0 caulobacter
18	458	21.4	895	2	Q61CU7_CABTR	Q61cu7 caenorhabdi
19	453.5	21.2	736	2	Q8A028_BACTN	Q8a028 bacteroides
20	449	21.0	745	2	Q6F317_PSED	Q6f317 pseudomonas
21	445.5	20.8	596	2	Q6K880_ORYSA	Q6k880 oryza sativ
22	442.5	20.7	745	2	Q5H5W8_XANTHOMAS	Q5h5w8 xanthomonas
23	442	20.7	552	2	Q8GUJ7_XANOR	Q8guj7 arabidopsis
24	442	20.7	746	2	Q9FNF6_ARATH	Q9fnf6 arabidopsis
25	441.5	20.6	757	2	Q8PFD7_XANAC	Q8pfd7 xanthomonas
26	440.5	20.6	766	2	Q4TNP1_9SPHN	Q4tnp1 erythrobact
27	436.5	20.4	763	2	Q8EAB7_SHEON	Q8eab7 shewanella
28	433.5	20.3	751	2	Q4UPD3_XANCP	Q4upd3 xanthomonas
29	433.5	20.3	751	2	Q8P3V8_XANCP	Q8p3v8 xanthomonas
30	432.5	20.2	743	2	Q5QX36_IDILO	Q5qx36 idiomarina
31	427.5	20.0	736	2	Q5LGU5_BACFN	Q5lgu5 bacteroides
32	427.5	20.0	736	2	Q64XP9_BACFR	Q64xp9 bacteroides
33	424	19.8	850	2	Q6H9E3_9TRYP	Q6h9e3 trypanosoma
34	409	19.1	852	2	Q4Q1H9_LEIMA	Q4q1h9 leishmania
35	400	18.7	720	2	Q5NMW8_ZYMMO	Q5nmw8 zymomonas m
36	398.5	18.6	741	2	P95782_XANMA	P95782 xanthomonas
37	393.5	18.4	732	2	Q7MUM6_PORGI	Q7mum6 porphyromon
38	383	17.9	749	2	Q7NEK8_GLOVI	Q7nek8 gloeobacter
39	379	17.7	711	2	Q47900_FLAME	Q47900 flavobacter
40	377.5	17.6	731	2	Q6L872_PREIN	Q6l872 prevotella
41	376	17.6	723	2	Q31048_PORGI	Q31048 porphyromon
42	376	17.6	723	2	Q66223_PORGI	Q66223 porphyromon
43	375	17.5	778	2	Q51ZP7_MAGRI	Q51zp7 magnaporthe
44	374.5	17.5	730	2	Q93JY4_9BACT	Q93jy4 prevotella
45	372	17.4	739	2	Q5FQY6_GLUOX	Q5fqy6 gluconobact

#### ALIGNMENTS

RESULT 1  
DPP8 HUMAN STANDARD: PRT: 898 AA.  
ID DPP8 HUMAN STANDARD: Q7ZAC8; Q7ZAE1; Q8TNG7; Q8NEM5; Q96JX1; Q9HEM2;  
AC Q6V1X1; Q7ZAC8; Q7ZAE1; Q8TNG7; Q8NEM5; Q96JX1; Q9HEM2;  
AC Q9HEM2; Q9HEM4; Q9HEM5; Q9NXP4;  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (dipeptidyl peptidase VIII) (DPP8)  
DE (Prolyl dipeptidase DPP8) (Dipeptidyl peptidase IV-related protein 1) (DPP8-1).  
DE Name=DPP8; Synonyms=DPRP1; ORFNames=MSTP097, MSTP135, MSTP141;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF 334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION, CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION, AND SUBCELLULAR LOCATION.  
RC TISSUE=Placenta;  
RX MEDLINE=20467194; PubMed=11012666;  
RA Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W., Gorrell M.D.;  
RT "Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8.";  
RL Eur. J. Biochem. 267:6140-6150(2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME REGULATION, BIOPHYSICOCHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
RC TISSUE=reticis;  
RX PubMed=12662155; DOI=10.1042/BJ20021914;  
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akinsanya K.O.;  
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member of an emerging subgroup of serine proteases.";  
RL Biochem. J. 373:179-189(2003).



[3] NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).  
RP TISSUE=Testis;  
RA Sha J.H., Zhou Z.M., Li J.M.;  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
[4]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4), AND NUCLEOTIDE  
RP SEQUENCE [LARGE SCALE MRNA] OF 211-898 (ISOFORM 2).  
RC TISSUE=Hepatoma, and placenta;  
EX PubMed=14702039, DOI=10.1038/ngl1285;  
RA Oka T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,  
RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,  
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Tanikawa M.,  
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,  
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,  
RA Fujinori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,  
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,  
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,  
RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,  
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakaagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakiabe H.,  
RA Hishigaki H., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
RA Yamazaki M., Watanabe K., Kumagai A., Tanigami S., Fukuzumi Y.,  
RA Fujinori Y., Komiyama M., Tashiro H., Tanigami A., Fujiiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Okamoto S.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Kawakami T., Noguchi S., Ichi T., Shigeoka K., Senba T.,  
RA Tatemura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Tsuchihashi-Sugano J., Saito T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakaki K., Ikeda I., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
RT "Complete sequencing and characterization of 21,243 full-length human  
RT cDNAs.";  
RL Nat. Genet. 36:40-45 (2004).  
[5]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).  
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whitling M., Helton E., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,  
RA Scherch A., Schin J.E., Jones S.J.M., Mirra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[6]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 561-898.  
RC TISSUE=Aorta;  
RA Zhao B., Xu H.S., Tong Y.K., Sheng H., Qin B.M., Liu Y.Q., Liu B.,  
RA Wang X.Y., Zhang Q., Song L., Gao Y., Zhang C.L., Ye J., Ji X.J.,  
RA Liu B.H., Lu H., Chen J.Z., Cai M.Q., Zheng W.Y., Teng C.Y., Liu Q.,  
RA Yu L.T., Lin J., Gong Q., Zhang A.M., Gao R.L., Hui R.T.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

[7] MUTAGENESIS OF GLU-275; SER-755; ASP-833 AND HIS-865, CATALYTIC  
RP ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, AND SUBCELLULAR LOCATION.  
RA PubMed=12534281; DOI=10.1021/b1026846s;  
RX Ajami K., Abbott C.A., Obradovic M., Gysbers V., Kaehne T.,  
RA McCaughan G.W., Gorrell M.D.;  
RT "Structural requirements for catalysis, expression, and dimerization  
RT in the CD26/DPPIV gene family.";  
RL Biochemistry 42:694-701 (2003).  
[8]  
RP BIOPHYSICOCHEMICAL PROPERTIES.  
RX PubMed=15039077; DOI=10.1016/j.pep.2003.12.019;  
RA Chen Y.-S., Chien C.-H., Goparaju C.M., Hsu J.T.-A., Liang P.-H.,  
RA Chen X.;  
RT "Purification and characterization of human prolyl dipeptidase DPP8 in  
RT Sf9 insect cells.";  
RL Protein Expr. Purif. 35:142-146 (2004).  
[9]  
RP ENZYME REGULATION.  
RX PubMed=15664838; DOI=10.1016/j.bmol.2004.11.023;  
RA Jiang W.-T., Chen Y.-S., Hsu T., Wu S.-H., Chien C.-H., Chang C.-N.,  
RA Chang S.-P., Lee S.-J., Chen X.;  
RT "Novel isoindoline compounds for potent and selective inhibition of  
RT prolyl dipeptidase DPP8.";  
RL Bioorg. Med. Lett. 15:687-691 (2005).  
CC -1- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal  
CC dipeptides from proteins having a Pro or Ala residue at position  
CC 2. May play a role in T-cell activation and immune function.  
CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-  
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided  
CC Zaa is neither Pro nor hydroxyproline.  
CC -1- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine  
CC proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride  
CC (AEBSF), and by di-isopropylfluorophosphate. Specifically inhibited  
CC by isoindoline derivatives.  
CC -1- BIOPHYSICOCHEMICAL PROPERTIES:  
CC Kinetic parameters:  
CC KM=208 uM for Ala-Pro-AMC;  
CC KM=130 uM for Ala-Pro-AFC;  
CC KM=120 uM for H-Ala-Pro-pNa;  
CC KM=1420 uM for H-Ala-Ala-pNa;  
CC KM=310 uM for H-Arg-Pro-pNa;  
CC KM=2050 uM for H-Asp-Pro-pNa;  
CC KM=480 uM for H-Gly-Pro-pNa;  
CC pH dependence:  
CC Optimum pH is 7.4-8.5. Little activity below pH 6.5;  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=6;  
CC Name=1:  
CC IsoId=Q6V1X1-1; Sequence=Displayed;  
CC Name=2:  
CC IsoId=Q6V1X1-2; Sequence=VSP\_013864;  
CC Name=3:  
CC IsoId=Q6V1X1-3; Sequence=VSP\_013860;  
CC Name=4:  
CC IsoId=Q6V1X1-4; Sequence=VSP\_013860; VSP\_013862;  
CC Name=5:  
CC IsoId=Q6V1X1-5; Sequence=VSP\_013863;  
CC Name=6:  
CC IsoId=Q6V1X1-6; Sequence=VSP\_013861;  
CC -1- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in  
CC testis, placenta, prostate, muscle and brain.  
CC -1- INDUCTION: In activated T-cells.  
CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.  
CC -1- CAUTION: Ref.4 (BAA91059) sequence differs from that shown due to  
CC a frameshift in position 486.  
CC -1- CAUTION: Ref.6 (AAQ13650 and AAQ13623) sequences differ from that  
CC shown due to several frameshifts.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not  
CC removed.

DR EMBL; AF221634; AAG29766.1; -; mRNA.  
DR EMBL; AF221635; AAG29767.1; -; mRNA.  
DR EMBL; AF221636; AAG29768.1; -; mRNA.  
DR EMBL; AF221637; AAG29769.1; -; mRNA.  
DR EMBL; AY172659; AAO17261.1; -; mRNA.  
DR EMBL; AY354202; AAQ3887.1; -; mRNA.  
DR EMBL; AK000290; BAA1059.1; ALT\_FRAME; mRNA.  
DR EMBL; AK027826; BAB55395.1; ALT\_FRAME; mRNA.  
DR EMBL; EC030688; AAH30688.3; -; mRNA.  
DR EMBL; EC040203; AAH40203.1; ALT\_INIT; mRNA.  
DR EMBL; AF176779; AAQ13657.1; ALT\_INIT; mRNA.  
DR EMBL; AF175225; AAQ13650.1; ALT\_FRAME; mRNA.  
DR EMBL; AF173382; AAQ13623.1; ALT\_FRAME; mRNA.

## Alignment Scores:

Pred. No.: 6,46e-144 Length: 898  
Score: 1645.50 Matches: 310  
Percent Similarity: 86.4% Conservative: 0  
Best Local Similarity: 86.4% Mismatches: 0  
Query Match: 76.9% Indels: 49  
DB: 1 Gaps: 1

US-10-825-632-4 (1-1197) x DPP8\_HUMAN (1-898)

QY 3 TTTGAGGACCAAGACTCCCTTTAGAGCATCCTCTAGCTAGTACGTTACGTAAAT 62  
DB PheGluGluThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 559  
QY 63 CTTGAGAGGTCACAGCGTACGACCGTGGCTACTCATTCTTGTGTCAGTCAG 122  
DB ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysLeuGln 579  
QY 123 CACTGTGACTCTTTTATAAGTATAGTATAGCAAGAAATCCACACTGTGTCCCTT 182  
DB HisCysAspPhePheLeuSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 599  
QY 183 TACAGCTATCAAGTCTGAGATGACCCCACTTGCAACAAAGAAATTTTGGCCACC 242  
DB TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 619  
QY 243 ATTTGGATTCCAGAGCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAA 302  
DB IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLeuPheSerPheGlu 639  
QY 303 AATCTACTGATTTACATTGATGGATGCTCTCAAGCCTCATGATCTACAGCTGGA 362  
DB SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 659  
QY 363 AAGAAATATCTACTGTGCTTCATATATATGTTGGTCTTCAG----- 404  
DB LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsn 679  
QY 404 ----- 404  
DB ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 699  
QY 404 ----- 404  
DB ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 719  
QY 405 -----GTCAAATAGAAATTCAGCATCAGTGGAGGACTCCATATCTAGCTTCT 455  
DB TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 739  
QY 456 CGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCAGCGTGTCTTATGGAGGATAC 515  
DB ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 759  
QY 516 CTCTCCCTGATGGCAATTAATGCAGAGTCAGATATCTTCAGGGTGTGCTATTCGTGGGCC 575

DB LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 779  
QY 576 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCT 635  
DB ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 799  
QY 636 GACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATCAAGCAGAAAAGTTCCCC 695  
DB AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlyLysPhePro 819  
QY 696 TCTGAACCAATCGTTTACTGCTTCTACATGTTCTCTGGATGAGAATGTCATTTCGCA 755  
DB SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 839  
QY 756 CATACCACTATATTACTGAGTGTCTTTAGTCAGGCTGGAAAGCCATATGATTTACAGATC 815  
DB HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 859  
QY 816 TATCTCAGAGGAGACACAGCATTAAGAGTTCTCGAATCGGAGAACATTAAGAACTGAT 875  
DB TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 879  
QY 876 CTTTGTGCACTACTCTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAAGTGATA 932  
DB LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 898  
RESULT 2  
DPP8\_MOUSE  
ID DPP8\_MOUSE STANDARD; PRT; 892 AA.  
AC Q80YA7; Q9D4G6;  
DT 13-SEP-2005 (Rel. 48, Created)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII)  
DE (D8)  
DE Name=Dpp8;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridea; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.W.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedziera R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sadelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sulcane R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang J., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume T., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."

Nature 420:563-573 (2002).  
 [2]  
 NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loughran N.A., Peters G.J., Abramson R.D., Mallaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Heiton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Whitney M., Madan A., Young A.C., Shevchenko K., Souffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Butcherfield Y.S.N., Krzywiniski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [3]  
 PHOSPHORYLATION SITE TYR-325.  
 RX PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;  
 RA Shu H., Chen S., Bi Q., Mumby M., Brekken D.L.;  
 RT "Identification of phosphoproteins and their phosphorylation sites in  
 the WEHI-231 B lymphoma cell line.";  
 RL Mol. Cell. Proteomics 3:279-286 (2004).  
 CC -1- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal  
 CC dipeptides from proteins having a Pro or Ala residue at position  
 CC 2. May play a role in T-cell activation and immune function (By  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-  
 CC Zaa- from a polypeptide, preferentially when Yaa is Pro, provided  
 CC Zaa is neither Pro nor hydroxyproline.  
 CC -1- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine  
 CC proteinase inhibitor 4-(2-aminocethyl)benzenesulphonyl fluoride  
 CC (AEBSP) and by di-isopropyl fluorophosphate. Specifically inhibited  
 CC by acidoaldehyde derivatives (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC EMBL; AK016546; BAB30295.2; -; mRNA.  
 CC EMBL; BC043124; AAH43124.1; -; mRNA.  
 CC EMBL; BC059222; AAH59222.1; -; mRNA.  
 CC MEROPS; S09.018; -.  
 CC DR Ensembl; ENSMUSG0000032393; Mus musculus.  
 CC DR MGI; MGI:1921638; Dpp8.  
 CC DR InterPro; IPR001375; Peptidase\_S9.  
 CC DR InterPro; IPR002469; Peptidase\_S9B.  
 CC DR InterPro; IPR000379; Ser\_estrs.  
 CC DR Pfam; PF00930; DPPIV\_N; 1.  
 CC DR Pfam; PF00326; Peptidase\_S9; 1.  
 CC DR AminoPeptidase; Hydrolase; Phosphorylation; Protease; Serine protease.  
 CC ACT SITE 749 749 Charge relay system (By similarity).  
 CC ACT SITE 827 827 Charge relay system (By similarity).  
 CC ACT SITE 859 859 Charge relay system (By similarity).  
 CC MOD\_RES 325 325 Phosphotyrosine.  
 CC FT CONFLICT 87 87 G -> R (in Ref. 1).  
 CC FT CONFLICT 87 87  
 CC SQ SEQUENCE 892 AA; 102186 MW; 59081CD9792E03ED CRC64;  
 Alignment Scores:  
 Pred. No.: 3.42e-140 Length: 892

Score: 1605.50 Matches: 302  
 Percent Similarity: 85.0% Conservative: 3  
 Best Local Similarity: 84.1% Mismatches: 5  
 Query Match: 75.0% Indels: 49  
 DB: 1 Gaps: 1  
 US-10-825-632-4 (1-1197) x DPP8\_MOUSE (1-892)  
 QY 3 TTTTGAAGGACCAAGAGATCCCTTTAGAGATCACCCTGTAGTACGTAGTACGTAGTACGTAAAT 62  
 DB 534 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValThrSerTyrAlaAsn 553  
 QY 63 CTGTGAGAGCGTCACAGGCGTACGACCGTGGCTACTCACAATCTTCTCTCATCAGTCAG 122  
 DB 554 ProGlyGluValValArgLeuThrAspArgGlyTyrSerHisSerCysCysLeuSerArg 573  
 QY 123 CACTGTGACTCTCTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCCTT 182  
 DB 574 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 593  
 QY 183 TACAAGCTATCAAGTCCTCTGAGATGACCCAACTTGTCAAAACAAAGGAATTTTGGGCCACC 242  
 DB 594 TyrLysLeuSerSerProGluAspAspProValHisLysThrLysGluPheTyrAlaThr 613  
 QY 243 ATTTTGGATTCAGCGGTCCTCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTTGA 302  
 DB 614 IleLeuAspSerAlaGlyProLeuProAspGlyThrProGluLysLysPheSerPheGlu 633  
 QY 303 AGTACTACTGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTTACAGCTCGGA 362  
 DB 634 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 653  
 QY 363 AAGAAATATCTACTGCTGCTTCATATATATGTTGGTCTCTCAG----- 404  
 DB 654 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 673  
 QY 404 ----- 404  
 DB 674 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 693  
 QY 404 ----- 404  
 DB 694 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 713  
 QY 405 -----GGTCAATAGAAATTCACCATCAGGTGGAGGACTCCAATATCTAGCTTCT 455  
 DB 714 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 733  
 QY 456 CGATATGATTTCATTCGATTCGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATAC 515  
 DB 734 GlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTyr 753  
 QY 516 CTCTCTCGATGCGATTAAATGACAGAGGTACAGATATCTTCAGGTTGCTATTGTGGGGCC 575  
 DB 754 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 773  
 QY 576 CCAGTCATCTGTGGATCTTCTATGATACAGGATACACGAGACGTTATATGGTGCACCT 635  
 DB 774 ProValThrLeuThrPheThrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 793  
 QY 636 GACCAAGATGAACAGGGCTATTACTAGATCTGTGGCATGCGAGGAGAGAAAGTCTCCC 695  
 DB 794 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 813  
 QY 696 TCTGAACCAAAATCGTTTACTGCTCTTATCATGTGTTCTCGATGAGATGTCATTTGCA 755  
 DB 814 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 833  
 QY 756 CATACCAAGTATATTACTAGTGTGTTTGTAGTGAGGCTGGAAAGCCATATGATTTTACAGATC 815  
 DB 834 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 853  
 QY 816 TATCTCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAATCGGGAGAACATATTGAACACTGCAT 875

Db 854 TyrProGInGluArgHisSerIleArgValProGluSerGlyCysHisTyrGluLeuHis 873  
QY 876 CTTTTCGCACTCCTTCAGAAACCTTGGATCAGCTATTGCTGCTCTAAAGTGATA 932  
Db 874 LeuLeuHisTyrLeuGInGluAsnLeuGlySerArgIleAlaAlaLeuHisValIle 892  
RESULT 3  
ID DPP9 MOUSE  
AC O8BVG4: O6RAM5; STANDARD; PRT; 862 AA.  
DT 13-SEP-2005 (Rel. 48, Created)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DP9)  
DE (Dipeptidyl peptidase-like protein 9) (DPLP9).  
GN Name=Dpp9;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridea; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
RC STRAIN=C57BL/6J; TISSUE=Liver, and Olfactory bulb;  
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard P., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wu L., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RT Nature 420:563-573 (2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).  
RC TISSUE=Thymus;  
RX PubMed=15449545;  
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,  
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.,  
RT "Prediction of the coding sequences of mouse homologues of FLJ genes:  
RT the complete nucleotide sequences of 110 mouse FLJ-homologous cDNAs  
RT identified by screening of terminal sequences of cDNA clones randomly  
RT sampled from size-fractionated libraries."  
RL DNA Res. 11:127-135 (2004).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
RC STRAIN=FVB/N; TISSUE=Salivary gland;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal  
CC dipeptides from proteins having a Pro or Ala residue at position 2  
CC (By similarity).  
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-  
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided  
CC Zaa is neither Pro nor hydroxyproline.  
CC -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-  
CC (2-aminoethyl)benzenesulphonyl fluoride (AEBSPF), and by di-  
CC isopropylfluorophosphate (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytosolic (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q8BVG4-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q8BVG4-2; Sequence=VSP\_013870, VSP\_013871, VSP\_013872;  
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.  
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a  
CC frameshift in position 745.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL: AK050021; BAC34034.1; -; mRNA.  
CC EMBL: AK078301; BAC37211.1; -; mRNA.  
CC EMBL: AK131178; BAD21428.1; ALT SEQ; mRNA.  
CC EMBL: BC057631; AAH57631.1; -; mRNA.  
CC MEROPS; S09.019; -; Mus musculus.  
CC Ensemble; ENSMUSG00000001229; Mus musculus.  
CC MGI: MGI:2443967; Dpp9.  
CC InterPro: IPR001375; Peptidase\_S9B.  
CC InterPro: IPR002469; Peptidase\_S9B.  
CC InterPro: IPR000379; Ser esters.  
CC Pfam: PF00930; DPPIV\_N; I.  
CC Pfam: PF00326; Peptidase\_S9; 1.  
CC Alternative splicing; Peptidase\_S9; 1.  
KW Serine protease.  
FT ACT\_SITE 729 729 Charge relay system (By similarity).  
FT ACT\_SITE 807 807 Charge relay system (By similarity).  
FT ACT\_SITE 839 839 Charge relay system (By similarity).  
FT VARSPPLIC 1 75 Missing (in isoform 2).  
FT FTID=VSP\_013870.  
FT VARSPPLIC 748 787 VAAGAPVTVMAYDTGYTYRMDVPENNQGYEAGSVAL  
FT -> PPEARSPSLSLPATDPDRMASASSSSWWEAKPGTASEG  
FT QR (in isoform 2).  
FT /FTID=VSP\_013871.  
FT Missing (in isoform 2).  
FT /FTID=VSP\_013872.  
FT D -> Y (in Ref. 1; BAC37211).  
FT CONFLICT 369 369 S -> Y (in Ref. 2).  
FT CONFLICT 546 546 S -> Y (in Ref. 2).  
FT CONFLICT 777 777 Q -> K (in Ref. 1; BAC37211).  
FT SEQUENCE 862 AA; 98001 MW; B1D566E824A834E8 CRC64;  
SQ

Alignment Scores:

Pred. No.: 1.3e-92 Length: 862  
Score: 1094.50 Matches: 200  
Percent Similarity: 69.3% Conservative: 42  
Best Local Similarity: 57.3% Mismatches: 58  
Query Match: 51.1% Indels: 49  
DB: 1 Gaps: 1

US-10-825-632-4 (1-1197) x DPP9\_MOUSE (1-862)

QY	3	TTTGAAGGACCAAGACCTCCCTTTAGAGATCACCTGTAGTCAGTACGTACGTAAAT	62
DB	514	PheGlnGlyThrIysAspThrProLeuGluHisLeuTyrValValSerTyrGluSer	533
QY	63	CTGGAGAGTGACAGGCTGACTGACCTGCTACACATCTTCTGCTCATCAGTCAG	122
DB	534	AlaGlyGluIleValArgGluThrThrLeuGlyPheSerHisSerCysSerMetSerGln	553
QY	123	CATGTGACTCTTTATAGTAAGTATAGTAACGAGAGATCCACACTGTGTCCCTT	182
DB	554	SerPheAspMetPheValSerHisTyrSerSerValSerThrProCysValHisVal	573
QY	183	TACAAGTATCAAGTCTCAAGATGACCCCACTTCCAAACAAAGGAATTTTGGCCACC	242
DB	574	TyrIysLeuSerGlyProAspAspProLeuHisIysGlnProArgPheThrPalaSer	593
QY	243	ATTGTGATTACAGCGTCTCTCTGACTACTCTCCAGAAATTTCTCTTTGAA	302
DB	594	MetMetGluAlaAlaAsnCysProProAspTyrValProGluIlePheHisPheHis	613
QY	303	AGTACTACTGGATTTACATTGTATGGATGCTCTCAAGCCTCATGATCTACAGCTGGA	362
DB	614	ThrArgAlaAspValGlnLeuTyrGlyMetIleTyrLysProHisThrLeuGlnProGly	633
QY	363	AAGAAATATCTACTGTCTGTTTCATATATATGTGTGCTCTCAG-----	404
DB	634	ArgIysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn	653
QY	404	-----	404
DB	654	SerPheIysGlyIleLysTyrLeuArgLeuLeuAsnThrLeuLeuSerLeuGlyTyrAlaVal	673
QY	404	-----	404
DB	674	ValValIleAspGlyArgGlySerCysGlnArgGlyLeuHisPheGluGlyAlaLeuLys	693
QY	405	-----GTCATATAGAAATGACGATCAGGTGGAGGACTCCATATCTAGCTTCT	455
DB	694	AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnTyrValAlaGlu	713
QY	456	CGATATGATTCATTGACTTAGATCGTGGGCATCCAGCGTCTCTATGGAGGATAC	515
DB	714	LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyIrpSerTyrGlyLysPhe	733
QY	516	CTCTCCCTGATGCAATTAATGACAGGTGAGATATCTCAGGCTTCTATGCTGGGCC	575
DB	734	LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla	753
QY	576	CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCA	635
DB	754	ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro	773
QY	636	GACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTCC	695
DB	774	GluAsnAsnGlnGlnGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro	793
QY	696	TCTGAACCAATCGTTTACTGCTCTTACATGTTCTCTGGATGAGAATGTCCATTTTGA	755
DB	794	AsnGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPhePhe	813
QY	756	CATACCAAGTATATTACTAGCTTTTATAGTGGGCTGGAAAGCCATATGATTTACAGATC	815
DB	814	HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle	833

QY	816	TATCTCTCAGGAGACACAGCATATAGATTCTCTGAATCGGAGACACATTATGACTGCAT	875
DB	834	TyrProAsnGluArgHisSerIleArgCysArgGluSerGlyGluHisTyrGluValThr	853
QY	876	CTTTTGCATCTTCAAGAAACCTT	902
DB	854	LeuLeuHisPheLeuGlnGluHisLeu	862
RESULT 4			
ID	DPP9_HUMAN	STANDARD;	PRT; 863 AA.
AC	Q86T12; Q6A137; Q6VAL0; Q6ZNT2; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WDX8;		
AC	Q96NT8; Q9BVR3;		
DC	13-SEP-2005 (Rel. 48, Created)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DE	Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)		
DE	Dipeptidyl peptidase-like protein 9) (DPLP9) (Dipeptidyl peptidase		
DE	IV-related protein 2) (DPRP-2).		
GN	Name=DPP9; Synonym=DPRP2;		
OS	Homo sapiens (human)		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.		
RP	MEDLINE=2347328; PubMed=12459266; DOI=10.1016/S0378-1119(02)01059-4;		
RA	Olsen C., Wegmann N.		
RT	"Identification and characterization of human DPP9, a novel homologue		
RT	of dipeptidyl peptidase IV."		
RL	Gene 299:185-193 (2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), CATALYTIC ACTIVITY.		
RP	BIOPHYSICO-CHEMICAL PROPERTIES, ENZYME REGULATION, TISSUE SPECIFICITY,		
RP	AND SUBCELLULAR LOCATION.		
RC	TISSUE=Colon;		
RX	PubMed=12662155; DOI=10.1042/BJ20021914;		
RA	Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akiansanya K.O.;		
RT	"Cloning and characterization of dipeptidyl peptidase 10, a new member		
RT	of an emerging subgroup of serine proteases."		
RL	Biochem. J. 373:179-189 (2003).		
RN	[3]		
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), PARTIAL NUCLEOTIDE SEQUENCE		
RP	[MRNA] (ISOFORM 2), CATALYTIC ACTIVITY, BIOPHYSICO-CHEMICAL PROPERTIES,		
RP	TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.		
RP	PubMed=15245913; DOI=10.1016/j.bbexp.2004.03.010;		
RA	Ajami K., Abbott C.A., McCaughan G.W., Gorrell M.D.;		
RT	"Dipeptidyl peptidase 9 has two forms, a broad tissue distribution,		
RT	cytoplasmic localization and DP1V-like peptidase activity."		
RL	Biochim. Biophys. Acta 1679:18-28 (2004).		
RN	[4]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).		
RC	TISSUE=Placenta; and Skin;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Wardeh R., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares W.B., Bonaldi M.F., Cabavant T.L., Schaeetz T.E.,		
RA	Brownstein M.J., Udwin T.B., Toshyuk S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters C.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richardson D.K., Wozny K.C., Hale S., Garcia A.M., Gibb R.A.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Hulton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting R.W., Touchman J.W., Green E.D., Bouffard G.G.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,		
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;		





Db 515 PheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGluAla 534  
Qy 63 CTGGAGAGTGACAAAGGCTGACTGACCGTGGCTACTACATTTCTTCTGCTGCATCAGTCAG 122  
Db 535 AlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGln 554  
Qy 123 CACTGTGACTCTTTTATAAGTATAGTAACACGAGAAGATCCACACTGTGTCCCTT 182  
Db 555 AsnPheAspMetPheValSerHisTyrSerSerValSerThrProProCysValHisVal 574  
Qy 183 TACAAGCTATCAAGTCTCAAGATGACCACTTGCACAAACAAAGGAAATTTGGGCCACC 242  
Db 575 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTetPAlaSer 594  
Qy 243 ATTTGGATTACAGCGTCTCTCTGACTATATCTCCAGAAATTTCTCTTTTGA 302  
Db 595 MetMetGluAlaAlaSerCysProProAspTyrValProGluIlePheHisPheHis 614  
Qy 303 AGTACTACTGGATTTACATTTGATGGGATGCTTACAGCGCTCATGATCTACAGCTGGA 362  
Db 615 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 634  
Qy 363 AAGAAATATCTACTGCTGCTTCATATATATGCTGCTCTCAG----- 404  
Db 635 LysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 654  
Qy 404 ----- 404  
Db 655 SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 674  
Qy 404 ----- 404  
Db 675 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 694  
Qy 405 -----GGTCAATAGAAATGACGATCAGGTGGAGGACCTCCAATATCTAGCTTCT 455  
Db 695 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 714  
Qy 456 CGATATGATTTCATTCAGCTTACATCGTGTGGGCATCCACGCGCTGCTCTATGGAGGATAC 515  
Db 715 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGlyPhe 734  
Qy 516 CTCTCCTGATGCTAATATGACAGAGTCCAGATCTTTCAGGGTTGCTATTGCTGGGCC 575  
Db 735 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 754  
Qy 576 CCACTCACTGTGGATCTTATGATACGATACAGGACGCTTATATGGTCCACCT 635  
Db 755 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 774  
Qy 636 GACCAAGATGACAGGCTATTACTTAGATCTGGCCATGCGAAGCAGAAAATTCCTCC 695  
Db 775 GluAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 794  
Qy 696 TCTGAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGATGTCATTTTGA 755  
Db 795 AsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePhe 814  
Qy 756 CATACCAATATATTACTAGTGTGTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATC 815  
Db 815 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 834  
Qy 816 TATCCTCAGGAGACACAGCATAAGATTCCTGAATCGGAGAACATATATGAAGTGCAT 875  
Db 835 TyrProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThr 854  
Qy 876 CTTTTCACCTACCTTCAAGAAACCTT 902  
Db 855 LeuLeuHisPheLeuGlnGluTyrLeu 863  
RESULT 5  
Q4SBM6\_TETNG  
ID Q4SBM6\_TETNG PRELIMINARY; PRT; 923 AA.

Q4SBM6; 2005 (T-EMBLrel. 31, Created)  
13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
13-SEP-2005 (T-EMBLrel. 31, Last annotation update)  
DE Chromosome 15 SCAF14667, whole genome shotgun sequence.  
DN (Fragment).  
DE ORFNames=GSTENG0020903001;  
OS Tetraodon nigroviridis (Green puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Tetraodon.  
OX NCBI\_Taxid=99883;  
RN [1]  
NUCLEOTIDE SEQUENCE.  
RP Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano B.,  
RA Anchoard V., Jubin C., Castellani V., Katinka M., Vacherie B.,  
RA Bismont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,  
RA Ceraud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,  
RA Parra G., Lardier S., Chappelle C., McKernan K.J., Mesirov J.,  
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Robinson-Rechavi M.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Scarpelli C.,  
RA Laudes V., Schachter V., Quetier P., Saurin W., Scarpelli C.,  
RA Winkler P., Landier E.S., Weissenbach J., Roest Crollius H.,  
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
the early vertebrate proto-karyotype.";  
RL Nature 431:946-957(2004).  
RN [2]  
NUCLEOTIDE SEQUENCE.  
RG Genoscope; Whitehead Institute Centre for Genome Research;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; CAAB01014667; CAG01956.1; -; Genomic\_DNA.  
FT NON\_TER 923 1  
FT NON\_TER 923 923  
SQ SEQUENCE 923 AA; 105211 MW; 1397023B2004D009 CRC64;  
  
Alignment Scores:  
Pred. No.: 1,13e-91 Length: 923  
Score: 1084.50 Matches: 199  
Percent Similarity: 66.6% Conservative: 52  
Best Local Similarity: 52.8% Mismatches: 49  
Query Match: 50.7% Indels: 77  
DB: 2 Gaps: 2  
  
US-10-825-632-4 (1-1197) x Q4SBM6\_TETNG (1-923)  
Qy 3 TTTCAGGACCAAGACCTCCCTTTAGGATCCTGCTAGCTAGTACGTTACGTAAT 62  
Db 547 PheGlnGlyThrArgAspThrProLeuGluHisHisLeuTyrValValSerTyrAspSer 566  
Qy 63 CTGGAGAGTGACAAAGGCTGACTGACCGTGGCTACTACATTTCTTCTGCTGCATCAGT--- 119  
Db 567 ProGlyAspValValArgLeuThrLysProGlyPheSerHisSerCysSerValSerGln 586  
Qy 119 ----- 119  
Db 587 ValLysLysSerLeuGlnSerAspTyrPhePheAsnTyrSerSerIleThrLeuProLeu 606  
Qy 120 -----CAGCACTGTGATCTCTTTTATAAGTAAGTAACTAGTAACACAG 158  
Db 607 SerLeuSerSerPheIleTrpGlnAsnPheAspPhePheValSerHisTyrSerVal 626  
Qy 159 AAGAATCCACATGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGGAAGATGACCCCACTTGC 218  
Db 627 CysThrProCysValHisValTyrLysLeuAsnSerSerGluSerAspProLeuHis 646  
Qy 219 AAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTTCTGACTATACT 278







```
Db 371 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 390
Qy 576 CcAGTCTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATATGGTCCACCT 635
Db 391 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluAaGlyTyrMetAspValPro 410
Qy 636 GACAGAAATGAACAGGGCTATTACTTAGATCTGTGGCATGCGAAGCAAGAAAGTTCCCC 695
Db 411 GluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 430
Qy 696 TcTCAACCAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAAATGTCATTTTGGCA 755
Db 431 AsnGlnProAsnArgLeuIleIleHisGlyPheLeuAspGluAsnValHisPhePhe 450
Qy 756 CATACCAATATATTACTGATGTTTATAGTACGGCTGGAAAGCCATATGATTTA----- 809
Db 451 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnVal 470
Qy 810 -----CAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTT 848
Db 471 AlaLeuProValSerProGlnIleTyrProAsnGluArgHisSerIleArgCysPro 490
Qy 849 GAATCGGGAGAACATATTGAATGCATCTTTTGGCACTACCTTCAAGAAACCTT 902
Db 491 GluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 508

RESULT 8
ID Q7PTT8_ANOGA PRELIMINARY; PRT; 621 AA.
AC Q7PTT8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE ENSANGP00000015447 (Fragment).
GN ORFNames-ENSANG00000012958;
OS Anopheles gambiae scr. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBSJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008287; EAA03335.3; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser esters.
DR Pfam; PF00930; DPPIV N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1
FT NON_TER 621
SQ SEQUENCE 621 AA; 70546 MW; F6EA8463A343BBF3 CRC64;

Alignment Scores:
Pred. No.: 9.3e-59 Length: 621
Score: 730.50 Matches: 154
Percent Similarity: 52.3% Conservative: 42
Best Local Similarity: 41.1% Mismatches: 92
```

```
Query Match: 34.1% Indels: 87
DB: 2 Gaps: 4
US-10-825-632-4 (1-1197) x Q7PTT8_ANOGA (1-621)
Qy 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTCTGCTAGTCTAGTTACGTAAAT 62
Db 258 PheMetGlyLeuArgGluThrProLeuGluLysHisLeuTyrValValSerLeuAlaGln 277
Qy 63 CTTGAGAGAGGTGACCAAGGCTGACTGACCGGTGGCTACTCATTCTTGTGTCATCAGTCAG 122
Db 278 ProAsnGlnLeuArgLeuLeuThrMetProGlyTyrSerPheThrValGluPheAsnAsp 297
Qy 123 CACTGTGACTCTTTTATAGTAAATAGTATAGTACCAAGAAATCCACACTGTGTGTCCTT 182
Db 298 AspCysThrLeuPheLeuGlnThrTyrCysAsnIleSerThrLeuProSerTrpGluLeu 317
Qy 193 TACAAGCTA-----TCAAGTCTCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 233
Db 318 ValArgIleAlaHisAspSerAsnThrAlaAsnGlyAsnGlyCys----- 332
Qy 234 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTACTACTCTCT----- 281
Db 333 -----SerHisGlyProThrProProThrProThrProIleAspAlaLeu 346
Qy 281 ----- 281
Db 347 ArgLeuCysSerValGlyTyrLeuThrGluGlyGlyProSerGluAsnThrGlnTyrAsn 366
Qy 282 CCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTTACATTCTATGGGATGCTCTACAG 341
Db 367 ProSerIleHisSerProGlnIleSerSerGlyAspValLeuTyrAlaMetValPheLys 386
Qy 342 CTTCTATGATCTACAGCCTGGAAGAAATATCTCTACTGTCTGTCTCATATATATGCTGCT 401
Db 387 ProHisAsnPheMetLeuGlyValLysTyrProThrValLeuAsnValTyrGlyGlyPro 406
Qy 402 CAG----- 404
Db 407 GluValGlnThrValSerAsnThrPheLysGlyMetArgGlnLeuArgMetHisMetLeu 426
Qy 404 ----- 404
Db 427 AlaSerGlnGlyTyrCysValIleCysValAspSerArgGlySerArgHisArgGlyVal 446
Qy 405 -----GGTCAAATAGAAATTCACGATCAGGTGCAA 434
Db 447 GluPheGluSerTyrIleArgArgArgMetGlyThrValGluLeuSerAspGlnValGlu 466
Qy 435 GGACTCCAATATCTAGTCTCTCGATATGATTTTCATTGACTTAGATCTGTGGGCACTCCAC 494
Db 467 ValLeuArgIleLeuAlaAspGlnLeuGlyTyrIleAspMetAspArgValAlaIleHis 486
Qy 495 GGCTGTGCTCTATGAGGATACCTCTCTCCCTGATGGCAATTAATCCAGAGGTGAGATCTTC 554
Db 487 GlyTrpSerTyrGlyGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProGluIlePhe 506
Qy 555 AGGGTGTCTATGCTGGGGCCCGCAGTCACTCTGTGATCTCTTATGATACAGGATACACG 614
Db 507 LysValSerIleAlaGlyAlaProValThrSerTrpGluTyrTyrAspThrGlyTyrThr 526
Qy 615 GAACGTTTATATGGGTCCACCTCGACCAAGTGAACAGGGCTATTACTAGGATCTCTGCCC 674
Db 527 GluArgTyrMetAspLeuProAspSerAsnArgSerGlyTyrAlaAlaGlySerValLeu 546
Qy 675 ATGCAAGCAAGAAAGTTCCCTCTCGAACCAATCTGTTACTCTCTTACATGCTTCTCTG 734
Db 547 AsnTyrIleGlnLysPheProAspGluAspAsnArgLeuIleIleHisGlyLeuLeu 566
Qy 735 GATGAGATGTCCATTTTTCACATACAGTATATATCTAGTCTTCTAGTGGGCTCGA 794
Db 567 AspGluAsnValHisPheHisThrSerGlnLeuValSerArgLeuValArgAlaAsn 586
```



OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;  
 OC Anophelinae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PEST;  
 RG The Anopheles gambiae Sequence Committee;  
 RT "Anopheles gambiae re-annotation."  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PEST;  
 RG The Anopheles gambiae Sequence Committee;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AA80100879; EAL40717.1; -; Genomic DNA.  
 DR GO; GO:0003824; F: catalytic activity; IEA.  
 DR GO; GO:0008236; F: serine-type peptidase activity; IEA.  
 DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001375; Peptidase\_S9.  
 DR InterPro; IPR002469; Peptidase\_S9B.  
 DR InterPro; IPR000379; Ser\_estrs.  
 DR Pfam; PF00930; DPPIV\_N; 1.  
 DR Pfam; PF00326; Peptidase\_S9; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 740 AA; 84151 MW; D59F131BBA095CCB CRC64;  
 Alignment Scores:  
 Pred. No.: 7,13e-57 Length: 740  
 Score: 710.50 Matches: 152  
 Percent Similarity: 54.4% Conservative: 44  
 Best Local Similarity: 42.2% Mismatches: 93  
 Query Match: 33.2% Indels: 71  
 DB: 2 Gaps: 4  
 US-10-825-632-4 (1-1197) x Q5T7K8\_ANOGA (1-740)  
 QY 3 TTGGAAGCCACCAAGACTCCCTTTAGAGCATACCTGTAGTAGTCAGTTACGTAAT 62  
 DB 392 PheMetGluLeuArgGluThrProLeuGluLysHisLeuTyrValValSerLeuAlaGln 411  
 QY 63 CTTGGAGGTCAGCAAGCGTCTGACCGTGGCTACTCAGCTTCTGCTGATCAGTCAG 122  
 DB 412 ProAsnGlnLeuArgLeuLeuThrMetProGlyTyrSer----- 424  
 QY 123 CACTGTGACTCTTTTATAGTAAGTATAGTAAC-----CAGAAGAATCCACAC 170  
 DB 425 -----PheThrValGluPheAsnAspValSerValPheGluArgLysGlyAla 440  
 QY 171 TGTGTGTCCTTTTACAGCTATCAGTCTCTGAGATGACCCAACTTGCAAAACAAGAA 230  
 DB 441 CysProAsnThrAlaAsnGlyAsnGlyCysSerHisGlyProThrProProThrPro 460  
 QY 231 TTTTGGGCGCCATTGTTGATTGATCAGCAGGTCTCTCTCTGACTATACCTCT- 281  
 DB 461 IleAspAlaLeuArgLeuCysSerValGlyTyrLeuThrGluGlyGlyProSerGluAsn 480  
 QY 282 -----CCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTCATCTGATGGG 329  
 DB 481 ThrGlnTyrAsnProSerIleHisSerProGlnIleSerSerGlyAspValLeuTyrAla 500  
 QY 330 ATGCTCTACAGCTCATGATCTACAGCTCTGGAAGAATATCTCTACTGCTGTTTCATA 389  
 DB 501 MetValPheLysProHisAsnPheMetMetLeuGlyValLysTyrProThrValLeuAsnVal 520  
 QY 390 TATGGTGGTCTCTCAG----- 404  
 DB 521 TyrGlyGlyProGluValGlnThrValSerAsnThrPheLysGlyMetArgGlnLeuArg 540  
 QY 404 ----- 404

Db 541 MethMetLeuAlaSerGlnGlyTyrCysValIleCysValAspSerArgGlySerArg 560  
 QY 405 -----GGTCAATAGAAATTCAC 422  
 DB 561 HisArgGlyValGluPheGluSerTyrIleArgCysArgMetGlyThrValGluLeuSer 580  
 QY 423 GATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATATTCAATGACTTACATCGT 482  
 DB 581 AspGlnValGluValLeuArgIleLeuAlaAspGlnLeuGlyTyrIleAspMetArg 600  
 QY 483 GTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAATGACAGAGG 542  
 DB 601 ValAlaIleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyr 620  
 QY 543 TCAGATATCTTCAGGGTGTATTGTCTGGGGCCCACTCTCTGTGATCTTCTTATGAT 602  
 DB 621 ProGluLeuPheLysValSerIleAlaGlyAlaProValThrSerTyrGluTyrTyrAsp 640  
 QY 603 ACAGATACACGGAACGTTATATGGTTCACCTGACCAAGATGAACAGGGCTATTACTTAA 662  
 DB 641 ThrGlyTyrThrGluArgTyrMetAspLeuProAspSerAsnArgSerGlyTyrAlaAla 660  
 QY 663 GGATCTGTGGCCATGCAAGCAGCAAAAGTCCCTCTGAACCAAAATCGTTACTGCTCTTA 722  
 DB 661 GlySerValLeuAsnTyrIleGlnLysPheProAspGluAspAsnArgLeuLeuIle 680  
 QY 723 CATGTTCTCTGGATGAGATGTCATTTGCACATACCACTATATTTACTGAGTTTTTTA 782  
 DB 681 HisGlyLeuIleAspGluAsnValHisPheHisHisThrSerGlnLeuValSerArgLeu 700  
 QY 783 GTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGA 842  
 DB 701 ValArgAlaAsnLysProTyrGlnLeuGlnValTyrProAsnGluArgHisSerLeuArg 720  
 QY 843 GTTCTCTGAATCGGGAGAACATTATGAACTGCACTCTTTTGCATCTACTCTTCAGAAAACCTT 902  
 DB 721 AsnLeuGluAlaSerLysHisTyrGluThrLysLeuLeuSerPheLeuGlnAsnHisLeu 740  
 RESULT 11  
 ID Q9VC20\_DROME PRELIMINARY; PRT; 1053 AA.  
 AC Q9VC20; Q8IH07;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
 DE CG3744-PA, isoform A (CG3744-pc, isoform c) (LD33755p).  
 GN ORFNames=CG3744;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cantarel D., Center A., Chandra I.,  
 RA Chervy J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Sridhan-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E.G., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yan R.-F., Zaveri J.S., Zhan N., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong E.W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:12185-12195 (2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celiniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirkas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
RT melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirkas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celiniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
RT a genome perspective.";  
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20 (2002).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.B., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celiniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Ruben S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review.";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RX Berkeley Drosophila Genome Project;  
RA Celiniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirkas R., Smith E.,  
RA Yu C., Rubin G.;  
RT "Drosophila melanogaster release 4 sequence.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RX FlyBase;  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokatein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dreanek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guatin H., Kronmiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celiniker S.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
CC -|- INTERACTION:  
CC Q9VZRO:CG12016; NbExp=1; IntAct=EBI-105926; EBI-152819;  
DR EMBL; AE003749; AAF56357.2; -; Genomic\_DNA.  
DR EMBL; BT001499; AAN71254.1; -; mRNA.  
DR IntAct; Q9VC20; -.  
DR MEROPS; S09.016; -.  
DR Ensembl; CG3744; Drosophila melanogaster.  
DR FlyBase; FBgn0039240; CG3744.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:0004374; F:dipeptidyl-peptidase IV activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001375; Peptidase\_S9.  
DR InterPro; IPR002469; Peptidase\_S9.  
DR InterPro; IPR000379; Ser. esters\_S9.  
DR Pfam; PF00930; DPPIV\_N; I.  
DR Pfam; PF00326; Peptidase\_S9; 1.  
SQ SEQUENCE 1053 AA, 118053 MW, C94AA663AB464577 CRC64;  
  
Alignment Scores:  
Pred. No.: 2,86e-56 Length: 1053  
Score: 704.50 Matches: 150  
Percent Similarity: 53.2% Conservative: 44  
Best Local Similarity: 41.1% Mismatches: 96  
Query Match: 32.9% Indels: 75  
DB: 2 Gaps: 6  
  
US-10-825-632-4 (1-1197) x Q9VC20\_DROME (1-1053)  
QY 3 TTTGAAGGCACCAAGACTCCCTTTAGACATACCTGTGACGTACGTACGTACGTACGTAAAT 62  
Db 694 PheValGlyLeuArgAspThrProLeuGluLysHisLeuTyValValSerLeuGluArg 713  
QY 63 CTGTGGAGAGTGCACAGGCTGACCGTGTGCTACTCCTACTCCTCTTCG----- 110  
Db 714 ProGluHisIleArgLeuLeuThrGluProGlyTySerTyTyValValGluPheAsp 733  
QY 111 -----TGCATCAGTCAGCACTGTGACTCTCTTTTATAAGTAAGTAT 149  
Db 734 HisPheAsnAspLeuPheProIlePheSerGlnGlnCysLysLeuMetLeuValTy 753  
QY 150 AGTAAC---CAGAAGAATCCACACTGTGTGCTCTTACAAAGCTATCAAGTCTCGAAGAT 206  
Db 754 CysAsnIleGlnArgLeuProSerCysLysValMet-----ArgVal 767  
QY 207 GACCAACTGTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTGACAGGT----- 260  
Db 768 AsnGlnThrCysSerAsnGlyGlyValAsnGlyIleGlnIleSerLeuValGlyTyLeu 787  
QY 261 -----CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGT 305  
Db 788 HisGluGlyGlyLysProGluProGlnTyCys---ProGlnIlePheSerProGlnLeu 806  
QY 306 ACTACTGGATTACATTGATGGATGCTCTACAGCTCATGATCTACAGCTGGGAAG 365  
Db 807 ProSerGlyAspValTyValMetValPheLysProHisAsnPheGluLeuGlyVal 826  
QY 366 AAATATCTACTGTCTTCTATATATGTGTCTCTCAG----- 404  
Db 827 LysTyProThrValLeuAsnValTyGlyGlyProGluValGlnThrValAsnAsnThr 846  
QY 404 ----- 404  
Db 847 PheLysGlyLysHisGlnLeuArgMetHisMetLeuAlaAlaGlnGlyTyCysValIle 866  
QY 404 ----- 404  
Db 867 CysIleAspSerArgGlySerArgHisArgGlyLysArgPheGluSerHisIleArgGly 886



```
QY 405 -----CGTCAATAGAAATGACGATCAGGTGGAGGACTCCCAATATCTAGCTTCGCA 458
Db 887 ArgMetGlyGlnValGluLeuThrAspGlnValAspAlaLeuArgSerAspGln 906
QY 459 TATGATTTCATGACTTACGATCTGTGGCCATCCAGCGCTGCTCTATGGAGGATACCTC 518
Db 907 LeuGlyTyrIleAspMetAspArgValAlaIleHisGlyTrpSerTyrGlyGlyTyrLeu 926
QY 519 TCCCTGATGCATTAATGACGAGGTCAGATATCTTTCAGGGTGTGCTATGCTGGGGCCCCA 578
Db 927 SerLeuMetGlyLeuValGlnTyrProLysIlePheLysValAlaIleAlaGlyAlaPro 946
QY 579 GTCACTCTGTGGATCTTCTATGATACAGATACAGGAACGCTATATATGGTCACCCCTGAC 638
Db 947 ValThrAsnTrpGluTyrAspThrGlyTyrThrGluArgTyrMetAspMetProGln 966
QY 639 CAGAAATCAACAGCGCTATTACTTATGATCTGTGGCCATGCAAGCAGAGAAAGCTCCCTCT 698
Db 967 AsnAsnGluAlaGlyTyrSerAlaGlySerValLeuGluTyrValAsnSerPheProGlu 986
QY 699 GAACCAATCGTTTACTGCTCTTACATGCTTCTCGATGAGATGTCCTCAATTTTGACAT 758
Db 987 GluAspLysArgLeuLeuLeuLeuHisGlyLeuIleAspGluAsnValHisPheCysHis 1006
QY 759 ACAGTATATATGAGTGTGTTTATGAGGCTGGAAAGCCATATGATTTACAGATCTAT 818
Db 1007 ThrSerArgLeuIleSerAlaLeuAsnLysAlaAsnLysProTyrGluValHisLeuPhe 1026
QY 819 CCTCAGAGAGACACGACATAGAGTCTCTGATCGGAGAACATTTAGACTGCATCTT 878
Db 1027 ProGluGluArgHisSerLeuArgAsnLeuGluSerAsnLysAsnTyrGluThrLysLeu 1046
QY 879 TTGCACCTACCTCCAA 893
Db 1047 LeuSerPheLeuGln 1051
RESULT 12
Q9VC19 DROME
ID Q9VC19 DROME PRELIMINARY; PRT: 1113 AA.
AC Q9VC19;
DT 01-MAY-2006 (TrEMBLrel. 13, Created)
DT 01-MAY-2006 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
CG3744-PE, isoform B.
GN Name=CG3744; ORFNames=CG3744;
OS Drosophila melanogaster (fruit fly);
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gilbert W.M., Glaeser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
```

```
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matti B., McIntosh T.C., McLéon M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bertram B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003749; AAF56356.2; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR Ensembl; CG3744; Drosophila melanogaster.
DR FlyBase; FBgn0039240; CG3744.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
```



```
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
SQ SEQUENCE 1113 AA; 123881 MW; 20857E3B212DF2E4 CRC64;

Alignment Scores:
Pred. No.: 2,9e-56 Length: 1113
Score: 704.50 Matches: 150
Percent Similarity: 53.2% Conservativeness: 44
Best Local Similarity: 41.1% Mismatches: 96
Query Match: 32.9% Indels: 75
DB: 2 Gaps: 6

US-10-825-632-4 (1-1197) x Q9VC19_DROME (1-1113)
QY 3 TTGGAAGGACCAAGACTCCCTTTAGACATCACCTGTAGTACGTAGTACGTAAAT 62
DB 754 PheValGlyLeuArgAspThrProLeuGluHisLeuTyrValValSerLeuGluArg 773
QY 63 CCTGGAGAGGTGACAGGCTGACTGACCGCTGGCTACTCACATCTTCTGC----- 110
DB 774 ProGluHisIleArgLeuLeuThrGluProGlyTyrSerTyrLeuValGluPheAspAsp 793
QY 111 -----TGCATCATGTCAGCAGCTGTGACTTCCTTTATAAGTAAGTAT 149
DB 794 HisPheAsnAspLeuPheProIlePheSerGlnGlnCysLeuMetLeuLeuValTyr 813
QY 150 AGTAAC---CAGAGATCCACACTGTGTGCTCCCTTTACAGCTATCAAGTCTCGAAT 206
DB 814 CysAsnIleGlnArgLeuProSerCysLeuValMet-----ArgVal 827
QY 207 GACCCAACTGCAAAACAAAGGAATTTGGGCGCCACTTTGGATTCAGCAGGT----- 260
DB 828 AsnGlnThrCysSerAsnGlyValAsnGlyLeGlnIleSerLeuValGlyTyrLeu 847
QY 261 -----CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGT 305
DB 848 HisGluGlyGlyLeuProGluProGlnTyrCys---ProGlnIlePheSerProGlnLeu 866
QY 306 ACTACTGGATTTACATTTGATGGATGCTCTACAGCTCATGATCTACAGCTCGAAAG 365
DB 867 ProSerGlyAspIleValTyrAlaMetValPheLeuProHisAsnPheGluLeuGlyVal 886
QY 366 AATATCTCTACTGTGTGCTTATATATATATGTTGCTCTCAG----- 404
DB 887 LysTyrProThrValLeuAsnValTyrGlyGlyProGluValGlnThrValAsnThr 906
QY 404 ----- 404
QY 907 PheLysGlyLysHisGlnLeuArgMetHisMetLeuAlaLaGlnGlyTyrCysValIle 926
QY 404 ----- 404
QY 927 CysIleAspSerArgGlySerArgHisArgGlyLysArgPheGluSerHisIleArgGly 946
QY 405 -----GGTCAATAGAAATTCACGATCAGGTGGAGGACTCCCAATATCTAGCTTCTCGA 458
QY 947 ArgMetGlyGlnValGluLeuThrAspGlnValAspAlaLeuArgSerLeuSerAspGln 966
QY 459 TATGATTTCAATGATAGTCTGTGGCATCCAGCTGGTCTCTATCGAGGATACCTC 518
DB 967 LeuGlyTyrIleAspMetAspArgValAlaIleHisGlyIlePheSerTyrGlyGlyTyrLeu 986
QY 519 TCCTGATGATGATTAATGACAGGTCAGATATCTTCAGGTTGCTATTGCTGGGGCCCA 578
DB 987 SerLeuMetGlyLeuValGlnTyrProLysIlePheLysValAlaIleAlaGlyAlaPro 1006
QY 579 GTCACTCTGTGGATCTTCTATGATACAGATACAGGAACTTATATGGGTCACTTCGAC 638
DB 1007 ValThrAsnTrpGlnTyrTyrAspThrGlyTyrGluArgTyrMetAspMetProGln 1026
QY 639 CAGAATGAACAGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAAAGATTCCTCTCT 698
```

---

```
Db 1027 AsnAsnGluAlaGlyTyrSerAlaGlySerValLeuGluTyrValAsnSerPheProGlu 1046
QY 699 GAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTCATTTGCAAT 759
Db 1047 GluAspLysArgLeuLeuLeuIleHisGlyLeuIleAspGluAsnValHisPheCysHis 1066
QY 759 ACCAGTATATTACTAGTATTTTATTAGTAGGGCTGGAAGCCATATGATTTACAGATCTAT 818
Db 1067 ThrSerArgLeuIleSerAlaLeuAsnLysAlaAsnLysProTyrGluValHisLeuPhe 1086
QY 819 CCTCAGGAGAGACACAGCATAAGAGTTCCTGAATCGGAGAACATTTAGAACTGCATCTT 878
Db 1087 ProGluGluArgHisSerLeuArgAsnLeuGluSerAsnLysAsnTyrGluThrLysLeu 1106
QY 879 TTGCACTACCTTCAA 893
Db 1107 LeuSerPheLeuGln 1111

RESULT 13
ID Q5TXJ2_ANOGA PRELIMINARY; PRT; 557 AA.
AC Q5TXJ2;
DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE ENSANGP0000029249 (Fragment).
CN ORFNames=ENSANG0000012958;
OS Anopheles gambiae str. PEST;
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008287; EAL42106.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1 557
FT NON_TER 557 557
SQ SEQUENCE 557 AA; 62685 MW; 08AAA0944ACD0808 CRC64;

Alignment Scores:
Pred. No.: 2,99e-50 Length: 557
Score: 639.00 Matches: 143
Percent Similarity: 53.9% Conservativeness: 35
Best Local Similarity: 43.3% Mismatches: 84
Query Match: 29.3% Indels: 68
DB: 2 Gaps: 6

US-10-825-632-4 (1-1197) x Q5TXJ2_ANOGA (1-557)
QY 3 TTGGAAGGACCAAGACTCCCTTTAGACATCACCTGTAGTACGTAGTACGTAAAT 62
Db 239 PheMetGlyLeuArgGlnThrProLeuGluLysHisLeuTyrValValSerLeuAlaGln 258
```



```
Db 576 Ty-Ser-Lys-Asp-Lys-Lys-Tyr-Pro-Thr-Val-Val-Tyr-Val-Tyr-Gly-Gly-Pro-His-Val-Gln 595
Qy 401 -----
Db 596 Ile-Val-Ile-Asn-Gln-Tyr-Asn-Tyr-Ile-Lys-Gln-His-Tyr-Thr-Asn-Phe-Gly-Phe-Ile-Gln 615
Qy 401 -----
Db 616 Val-Met-Ile-Asp-Asn-Val-Gly-Ser-Ala-Asn-Arg-Gly-Leu-Glu-Phe-Glu-Ser-His-Ile-Arg 635
Qy 402 -----
Db 636 Glu-Lys-Met-Gly-Gln-Val-Glu-Ile-Gly-Asp-Gln-Val-Glu-Gly-Ile-Asn-Tyr-Leu-Val-Gly 655
Qy 456 CGATATGATTTC-----ATTGACATTAGATCGTGTGGGCATCCACGGCTGCTCTATGGA 509
Db 656 Asn---Asp-Ile-Val-Ser-Ile-Asp-Val-Asn-Arg-Ile-Ala-Ile-Ser-Gly-Tyr-Ser-Tyr-Gly 674
Qy 510 GGATACCTCTCCCTGATGCATTAAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCT 569
Db 675 Gly-Tyr-Asn-Ser-Leu-Met-Ala-Ile-Ser-Gln-Arg-Pro-Asp-Val-Phe-Lys-Ile-Ala-Val-Cys 694
Qy 570 GGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACCTGATATGGGT 629
Db 695 Gly-Val-Pro-Val-Ser-Asp-Tyr-Arg-Leu-Tyr-Asn-Thr-Gly-Tyr-Thr-Glu-Arg-Tyr-Met-Asn 714
Qy 630 CACCTGACACAGGATGACAGGCTATTACTTACTAGATCTGTGGCATCTGCGCATCGACAGCAAAAG 689
Db 715 Val-Pro-Gln-Asp-Asn-Ile-Asp-Gly-Tyr-Lys-Leu-Gly-Asp-Thr-Thr-His-Tyr-Ser----- 732
Qy 690 TTCCCTCTGAAACCAATCGTTTACTCTGCTTACATGTTTCTGGAATGAGATGTCCTAT 749
Db 733 Phe-Pro-Thr-Glu-Glu-Asn-Arg-Leu-Leu-Ile-His-Gly-Leu-Gln-Asp-Glu-Asn-Val-His 752
Qy 750 TTGTCATACACAGTATATTACTAGTGTGTTTGTAGAGGGCTGGAAGCCATATGATTTA 809
Db 753 Phe-Ser-Asn-Thr-Ile-Tyr-Ile-Ile-Asp-His-Leu-Thr-Lys-Thr-Gln-Lys-Pro-Tyr-Ile-Leu 772
Qy 810 CAGATCTATCTCAGAGAGACACAGCATACAGATGTTCTGAAATCGGGAGAACATTATGAA 869
Db 773 Lys-Thr-Leu-Pro-Asn-Glu-Arg-His-Gly-Val-Arg-Asn-Thr-Asp-Asn-Arg-Ile-Tyr-Ile-Gly 792
Qy 870 CTGCATCTTTGTCACCTACCTTCAAGAAACCTT 902
Db 793 Leu-Phe-Val-Ile-Asn-His-Leu-Leu-Lys-Asn-Leu 803

RESULT 15
Q965K3 CABEL PRELIMINARY; PRT; 927 AA.
AC Q965K3 CABEL PRELIMINARY; PRT; 927 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Di-peptidyl peptidase four (iv) family protein 3, isoform b.
GN Name=dpf-3; ORFNames=K02F2.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology.";
RL Science 282:2012-2018 (1998).
DR EMBL; AF043699; AAK84627.1; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR Uniprot; K02F2.1; Caenorhabditis elegans.
DR WormBase; WBGene0001056; dpf-3.
DR WormPep; K02F2.1b; CE28837.
DR GO; GO:0016020; C:membrane; IEA.
```

```
DR GO; GO:0003244; F:catalytic activity; IEA.
DR GO; GO:0004274; F:di-peptidyl-peptidase IV activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR Complete proteome.
SQ SEQUENCE 927 AA; 105366 MW; 6907C6AAEF829D0F CRC64;

Alignment Scores:
Pred. No.: 1.98e-37 Length: 927
Score: 502.00 Matches: 122
Percent Similarity: 48.4% Conservative: 56
Best Local Similarity: 33.2% Mismatches: 104
Query Match: 23.5% Indels: 86
DB: Gaps: 14

US-10-825-632-4 (1-1197) x Q965K3 CABEL (1-927)
Qy 3 TTGTGAAGGCCAACAACTCCCTTTAGAGCATCACCTGACGTAGTACGTACGTAAAT 62
Db 559 Tyr-Val-Ala-Asn-Glu-Ser-His-Pro-Thr-Glu-Tyr-Phe-Ile---Cys-Val-Ser-His-Tyr-Arg 577
Qy 63 CCTGGAGAGGTGACAGAGGCTGACTGACCTGCGTACTCACATCTCTGCTGCATCAGTCAG 122
Db 578 Thr-Gly-Gln-His-Ala-Gln-Leu-Thr-Glu-Ser-Gly-Ile----- 589
Qy 123 CACTGTGACTCTTTTATAAGTAAGTATAGTATACCAAGAG-----AATCCACAC 170
Db 590 ---Cys-----Phe-Lys-Ser-Glu-Arg-Ala-Asn-Gly-Lys-Leu-Ala-Leu-Asp-Lys-Asp-His 606
Qy 171 TGTGTGTCCCTTTACAGCTATCA-----AGTCTGAAGATGACCCCACTTTGCAAAACA 224
Db 607 Gly-Phe-Ala-Cys-Tyr-Met-Thr-Ser-Val-Gly-Ser-Pro-Ala-Glu-----Cys-Arg-Phe 623
Qy 225 AAGCAATTT---TCG-----GCGACCAT 245
Db 624 Tyr-Ser-Phe-Arg-Tyr-Lys-Glu-Asn-Glu-Val-Leu-Pro-Ser-Thr-Val-Tyr-Ala-Ala-Ser-Ile 643
Qy 246 TTGGATTTCAGCAGCTCCT---CTTCTCTGAC-----TATCTCTCTCCAGAAATTTCTCT 296
Db 644 Thr-Val-Ser-Gly-His-Pro-Gly-Gln-Pro-Asp-Leu-His-Phe-Asp-Ser-Pro-Glu-Met-Ile-Glu 663
Qy 297 TTTGAAAGT---ACTACTGGATTTCATATTGATGGGATGCTCTACACACCTCATGATCTA 353
Db 664 Phe-Gln-Ser-Lys-Lys-Thr-Gly-Leu-Met-His-Tyr-Ala-Met-Ile-Leu-Arg-Pro-Ser-Asn-Phe 683
Qy 354 CAGCTCGAAAGAAATATCTCTGCTGTCTCATATATATGTTGGTGTCT----- 401
Db 684 Asp-Pro-Tyr-Lys-Lys-Tyr-Pro-Val-Phe-His-Tyr-Val-Tyr-Gly-Gly-Pro-Gly-Ile-Gln-Ile 703
Qy 401 ----- 401
Db 704 Val-His-Asn-Asp-Phe-Ser-Trp-Ile-Gln-Tyr-Ile-Arg-Phe-Cys-Arg-Leu-Gly-Tyr-Val-Val 723
Qy 401 ----- 401
Db 724 Val-Phe-Ile-Asp-Asn-Arg-Gly-Ser-Ala-His-Arg-Gly-Ile-Glu-Phe-Glu-Arg-His-Ile-His 743
Qy 402 -----CAGGCTCAAATAGAAATTGCGATCGCTGGAGGCTCCATATCTAGTCTCT 455
Db 744 Lys-Lys-Met-Gly-Thr-Val-Glu-Glu-Val-Glu-Asp-Gln-Val-Glu-Gly-Leu-Gln-Met-Leu-Ala-Glu 763
Qy 456 CGATAT---GATTTCATTGACTTAGATCTGTGGGCATCCACGGCTGCTCTATGAGGA 512
Db 764 Arg-Thr-Gly-Gly-Phe-Met-Asp-Met-Ser-Arg-Val-Val-His-Gly-Tyr-Ser-Tyr-Gly-Gly 783
Qy 513 TACCTCTCTCTGATGTCATTAAATGACAGAGTCCAGATATCTTCAGGGTTGCTATTGCTGGG 572
Db 784 Tyr-Met-Ala-Leu-Gln-Met-Ile-Ala-Lys-His-Pro-Asn-Ile-Tyr-Arg-Ala-Ile-Ala-Gly 803
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 01:34:19 ; Search time 7.33538 Seconds  
(without alignments)  
2698.234 Million cell updates/sec

Title: US-10-825-632-4  
Perfect score: 2140  
Sequence: 1 atttgaaggaccacaaagac.....ttactataaaaaaaaaa l197

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlp  
-Q/abss/ABSSWEB.spool/US10825632/runat\_01052006\_105950\_3311/app\_query.fasta\_1  
-DB=Issued\_Patents\_AA -QMT=fastan -SUFFIX=n2p.ra -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs02p -USER=US10825632 @CGN 1.1.131 @runat\_01052006\_105950\_3311  
-NCPU=6 -ICPU=3 -NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5 COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/iaa/6 COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/iaa/H COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/iaa/RE COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1680	78.5	310	2	US-09-794-236-4
2	1680	78.5	310	2	US-10-070-464-3
3	1645.5	76.9	882	2	US-09-976-674-1
4	1645.5	76.9	882	2	US-10-070-464-1
5	1377.5	64.4	465	2	US-10-070-464-5
6	1093.5	51.1	863	2	US-09-976-674-3
7	1093.5	51.1	892	2	US-09-976-674-23
8	1093.5	51.1	892	2	US-09-976-674-27
9	1086.5	50.8	360	2	US-10-070-464-7
10	1086.5	50.8	879	2	US-09-976-674-33
11	1086.5	50.8	879	2	US-09-976-674-35
12	867	40.5	690	2	US-09-976-674-7

13	739	34.5	658	2	US-09-976-674-19	Sequence 19, Appl
14	739	34.5	661	2	US-09-976-674-11	Sequence 11, Appl
15	629.5	29.4	832	2	US-09-976-674-29	Sequence 29, Appl
16	629.5	29.4	832	2	US-09-976-674-31	Sequence 31, Appl
17	622.5	29.1	819	2	US-09-976-674-37	Sequence 37, Appl
18	622.5	29.1	819	2	US-09-976-674-39	Sequence 39, Appl
19	471	22.0	613	2	US-09-976-674-21	Sequence 21, Appl
20	397	18.6	981	2	US-09-902-540-16812	Sequence 16812, A
21	393.5	18.4	732	2	US-09-518-550-30	Sequence 30, Appl
22	376	17.6	676	2	US-09-518-550-42	Sequence 42, Appl
23	376	17.6	723	2	US-09-518-550-29	Sequence 29, Appl
24	375	17.5	710	2	US-09-518-550-28	Sequence 28, Appl
25	337.5	15.8	766	2	US-10-002-593-6	Sequence 6, Appl
26	337.5	15.8	766	2	US-09-949-016-6146	Sequence 6146, Ap
27	337.5	15.8	766	2	US-10-423-7114-6	Sequence 6, Appl
28	333.5	15.6	755	4	PCT-US93-07923-3	Sequence 3, Appl
29	333.5	15.6	759	4	PCT-US93-07923-2	Sequence 2, Appl
30	333.5	15.6	766	1	US-08-230-491A-3	Sequence 3, Appl
31	333.5	15.6	766	1	US-08-619-280A-3	Sequence 3, Appl
32	333.5	15.6	766	1	US-08-940-391-3	Sequence 3, Appl
33	333.5	15.6	766	2	US-09-794-236-1	Sequence 1, Appl
34	333.5	15.6	766	2	US-09-265-606-3	Sequence 3, Appl
35	333.5	15.6	775	2	US-09-949-016-10450	Sequence 10450, A
36	329.5	15.4	766	2	US-09-518-550-27	Sequence 27, Appl
37	327	15.3	771	2	US-09-462-284-2	Sequence 2, Appl
38	326.5	15.3	737	2	US-09-079-592-2	Sequence 2, Appl
39	326.5	15.3	737	2	US-09-502-540-11421	Sequence 11421, A
40	317.5	14.8	323	2	US-09-270-767-45296	Sequence 45296, A
41	314	14.7	931	2	US-09-079-592-11	Sequence 11, Appl
42	310	14.5	818	2	US-09-462-845-3	Sequence 3, Appl
43	310	14.5	818	2	US-10-402-312-3	Sequence 3, Appl
44	310	14.5	818	2	US-10-401-437-3	Sequence 3, Appl
45	310	14.5	818	2	US-10-402-067-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-09-794-236-4  
; Sequence 4, Application US/09794236  
; Patent No. 6337069  
; GENERAL INFORMATION:  
; APPLICANT: Lacroix, Jean-Silvain  
; APPLICANT: Monod, Michel  
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis  
; FILE REFERENCE: 81985/276823  
; CURRENT FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 310  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-794-236-4

Alignment Scores:  
Pred. No.: 7.93e-191 Length: 310  
Score: 1680.00 Matches: 310  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 78.5% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-4 (1-1197) x US-09-794-236-4 (1-310)

QY	3	TTTGAAGCACAAGACTCCCTTTAGACATCACTGCTAGTGTAGTACGTTACGTAAT 62
Db	1	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 20
QY	63	CCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCATCTTGTGTCATGCTAGT 122

```
Db 21 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 40
Qy 123 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTT 182
Db 41 HisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSerLeu 60
Qy 183 TACAAGCTATCAAGTCTCAAGTACACCAACTTGCAGAAACAAAGGAATTTTGGGCACC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrPalaThr 80
Qy 243 ATTTTGGATTACAGAGTCTCTCTCTCACTATATCTCTCCAGAAATTTTCTCTTTGAA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 100
Qy 303 AGTACTACTGGATTTACATTGATGGATGCTCTACAGGCTCATGATCTACAGCTGGA 362
Db 101 SerThrThrGlyPheThrLeuTyrGlyMetLysThrLysGluPheTyrPalaThr 120
Qy 483 GTGGGCATCCACGGTGTCTATGAGGATACCTCTCCCTGACCAATTCCTTTACTGCTCTTA 722
Db 201 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAenGluGlnGlyTyrTyrLeu 220
Qy 663 GGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATTCCTTTACTGCTCTTA 782
Db 241 HisGlyPheLeuAspGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
Qy 783 GTGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAAGA 842
Db 261 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 280
Qy 843 GTTCCTGAATCGGAGAACATTATGAAGTGCATCTTTTGCACCTACCTTCAAGAAACCTT 902
Db 281 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAenLeu 300
Qy 903 GGATCAGTATTCCTCTCTAAAGTGATA 932
Db 301 GlySerArgIleAlaAlaLeuLysValIle 310
```

## RESULT 2

```
US-10-070-464-3
; Sequence 3, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
```

```
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-3
```

```
Alignment Scores:
Pred. No.: 7,93e-191 Length: 310
Score: 1680.00 Matches: 310
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 78.5% Indels: 0
DB: Gaps: 0
```

US-10-825-632-4 (1-1197) x US-10-070-464-3 (1-310)

```
Qy 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGTTACGTAAT 62
Db 1 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAen 20
Qy 63 CCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCTTCTGCTCAGTCAG 122
Db 21 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 40
Qy 123 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATTTTGGGCACC 182
Db 41 HisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSerLeu 60
Qy 183 TACAAGCTATCAAGTCTCAAGTACACCAACTTGCAGAAACAAAGGAATTTTGGGCACC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrPalaThr 80
Qy 243 ATTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 100
Qy 303 AGTACTACTGGATTTTACATTGATGGATGCTCTACAGCCTCATGATCTACAGCTGGA 362
Db 101 SerThrThrGlyPheThrLeuTyrGlyMetLysThrLysGluPheTyrPalaThr 120
Qy 363 AAGAAATATCCTACTGTGTCTTATATATGTTGGTGTCTCAGGTCATAAGAAATTTGAC 422
Db 121 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnGlyGlnIleGluIleAsp 140
Qy 423 GATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGT 482
Db 141 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 160
Qy 483 GTGGGCATCCACGGCTGTCTCTATGAGGATACCTCTCCCTGATGGCATTAATTCAGAGG 542
Db 161 ValGlyIleHisGlyTyrSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 180
Qy 543 TCAGATATCTTCAGGGTGTCTATTGTGGGGCCCGCAGTCCTCTGTGGATCTTCTATGAT 602
Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTyrPheTyrAsp 200
Qy 603 ACAGGATACACGGAACGTTATATGGTTCACCTCAGCAGAGAAATGAACGGGCTATTACTTA 662
Db 201 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAenGluGlnGlyTyrTyrLeu 220
Qy 663 GGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATTCCTTTACTGCTCTTA 722
Db 221 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAenArgLeuLeuLeu 240
Qy 723 CATGGTTTCTCGATGAGATCTCCATTTTGCACATACCATGATATATTTACTCAGTTTATA 782
Db 241 HisGlyPheLeuAspGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
Qy 783 GTGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAAGA 842
```



Db 261 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 280  
Qy 843 GTTCCTGAATCGGAGAACATTATCAACTGATCTTTTGGCACTACCTTCAAGAAACCTT 902  
Db 281 ValProGluSerGlyGluHisTyrGluLeuHisLeuHisLeuHisTyrLeuGlnGluAsnLeu 300  
Qy 903 GGATCAGTATTGCTGCTCTAAAGTGATA 932  
Db 301 GlySerArgIleAlaLeuLysValIle 310  
RESULT 3  
US-09-976-674-1  
; Sequence 1, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-1  
Alignment Scores:  
Pred. No.: 1,84e-186 Length: 882  
Score: 1645.50 Matches: 310  
Percent Similarity: 86.4% Conservative: 0  
Best Local Similarity: 86.4% Mismatches: 0  
Query Match: 76.9% Indels: 49  
DB: 2 Gaps: 1  
US-10-825-632-4 (1-1197) x US-09-976-674-1 (1-882)  
Qy 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATACCTGTAGTACGTAGTCAGTAAAT 62  
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543  
Qy 63 CTTGGAGAGTGACCAAGCTGACTGACCGTGCTACTACATTTCTGCTGATCAGTCAG 122  
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563  
Qy 123 CACTGTGACTTCTTTATAGTAACTAGTATAGTACAGAGATCCACACTGTGTCCCTT 182  
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
Qy 183 TACAAGCTATCAAGTCTCAAGATCACCAACTTGCACCAACAAAGGAATTTGGGCACC 242  
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
Qy 243 ATTTGGATTACAGAGGCTCTTCTGCTGATATCTCTCAGAAATTTCTTTTGAA 302  
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
Qy 303 AGTACTACTGGATTACATTTGATGATGCTTACAGCTCATGATCTACAGCTTGA 362  
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
Qy 363 AAGAAATATCTACTGTGCTTATATATATATATATATATATATATATATATATATATAT 404  
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn 663  
Qy 404 ----- 404

Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
Qy 404 ----- 404  
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
Qy 405 -----GGTCAAAATAGAAATTCAGATCAGGTGGAGGACTCCAATATCTAGCTTCT 455  
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
Qy 456 CGATATGATTTTCACTTACTAGATCGTGTGGGCATCCAGCGCTGCTATATGGAGGATAC 515  
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyrGlyGlyTyr 743  
Qy 516 CTCTCCCTGATGCGATTAAATGCAGAGGTGAGATATCTTCAGGGTGTCTATTGCTGGGGCC 575  
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
Qy 576 CCAGTCACTCTGTGGATCTTCTATATGATACAGGATACAGGAACGTTATATGGTCCACCT 635  
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783  
Qy 636 GACCAGATGACACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAGCAAGAAAGTTCCCC 695  
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803  
Qy 696 TCTGAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGAATGTCATTGTCAT 755  
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823  
Qy 756 CATAACAGTATATTACTAGTATTTTGTAGTGGCTGGAAAGCCATATGATTTACAGATC 815  
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843  
Qy 816 TATCTCAGAGACACACAGCATATAGATTTCTGATCGGAGACATATGAACTGCAT 875  
Db 844 TyrProGlnArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863  
Qy 876 CTTTGTGACTACCTTCAAGAAACCTTGGATCACGCTATTGCTGCTCTCTAAAAGTGATA 932  
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882  
RESULT 4  
US-10-070-464-1  
; Sequence 1, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-070-464-1  
Alignment Scores:  
Pred. No.: 1,84e-186 Length: 882  
Score: 1645.50 Matches: 310  
Percent Similarity: 86.4% Conservative: 0  
Best Local Similarity: 86.4% Mismatches: 0

Query Match: 76.9% Indels: 49  
DB: 2 Gaps: 1  
US-10-825-632-4 (1-1197) x US-10-070-464-1 (1-882)  
Qy 3 TTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACCTGACGAGTACGTTACGTTAAAT 62  
Db PheGluGlyThrLysAspSerProLeuGluHisLeuTyValValSerTyValAsn 543  
Qy 63 CTGGAGAGGTCAAGGCTGACTGACCGTGGCTACTCACATCTTCTGCTCATCAGTCAG 122  
Db ProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCysCysIleSerGln 563  
Qy 123 CACTGTGACTTCTTTATAGTAAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTT 182  
Db HisCysAspPhePheIleSerLysTySerAsnGlnLysAsnProHisCysValSerLeu 583  
Qy 183 TACAAGCTATCAAGTCTGAGAGTACCAACTTGGCAAAACAAGAGGAATTTTGGGCCACC 242  
Db TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrPalaThr 603  
Qy 243 ATTTTGGATTACAGCGTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAA 302  
Db IleLeuAspSerAlaGlyProLeuProAspTyThrProGluIlePheSerPheGlu 623  
Qy 303 AGTACTATCGGATTTACATTTGATGATGATGATGATGATGATGATGATGATGATGATG 362  
Db SerThrGlyPheThrLeuTyGlyMetLeuTyLysProHisAspLeuGlnProGly 643  
Qy 363 AGAATATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404  
Db LysLysTyProThrValLeuPheIleTyGlyProGlnValGlnLeuValAsn 663  
Qy 404 ----- 404  
Db 664 ArgPheLysGlyValLysTyPheArgLeuAsnThrLeuAlaSerLeuGlyTyValVal 683  
Qy 404 ----- 404  
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
Qy 405 -----GGTCAAATAGAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 455  
Db 704 TyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyLeuAlaSer 723  
Qy 456 CGATATGATTTCATTGACTTAGATCGTGTGGGCATCCACGCTGGTCTCTATGGAGGATAC 515  
Db 724 ArgTyAspPheIleAspLeuAspArgValGlyIleHisGlyTyPheSerTyGlyGlyTy 743  
Qy 516 CTCTCCCTGATGSCATTAATGCAGAGTACAGATATCTTCAGGTTGCTATTGCTGGGCC 575  
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
Qy 576 CCAGTCACCTCTGGATCTTCTATGATACAGGATACACGGAAGCTTATATGGTCACTCC 635  
Db 764 ProValThrLeuThrPheTyAspThrGlyTyThrGluArgTyMetGlyHisPro 783  
Qy 636 GACCAAGATGAACAGGCTATTACTAGGATCTGTGGCCATGACAGCAAGAAATGTCCTCC 695  
Db 784 AspGlnAsnGlnGlnGlyTy 803  
Qy 696 TCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGATGAGAAATGTCATTTTGA 755  
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823  
Qy 756 CATACAGPATATTACTAGTTTTTTAGTGAGGCTCGAAAGCCATATGATTTACAGATC 815  
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyAspLeuGlnIle 843  
Qy 816 TATCCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTAATGAACGTGAT 875  
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyGluLeuHis 863

Qy 876 CTTTTCACACTACCTTCAAGAAAACCTTGGAYCACGTATTGCTGCTCTAAAAGTGATA 932  
Db 864 LeuLeuHisTyLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882  
RESULT 5  
US-10-070-464-5  
; Sequence 5, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH 465  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-070-464-5  
Alignment Scores:  
Pred. No.: 1,08e-154 Length: 465  
Score: 1377.50 Matches: 259  
Percent Similarity: 83.5% Conservative: 0  
Best Local Similarity: 83.5% Mismatches: 0  
Query Match: 64.4% Indels: 51  
DB: 2 Gaps: 1  
US-10-825-632-4 (1-1197) x US-10-070-464-5 (1-465)

Qy 3 TTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACCTGATAGTACGTTACGTTAAAT 62  
Db 207 PheGluGlyThrLysAspSerProLeuGluHisLeuTyValValSerTyValAsn 226  
Qy 63 CTGGAGAGGTCAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTCATCAGTCAG 122  
Db 227 ProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCysCysIleSerGln 246  
Qy 123 CACTGTGACTTCTTTATAGTAAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTT 182  
Db 247 HisCysAspPhePheIleSerLysTySerAsnGlnLysAsnProHisCysValSerLeu 266  
Qy 183 TACAAGCTATCAAGTCTGAGAGTACCAACTTGCRAAACAAAGGAATTTTGGGCCACC 242  
Db 267 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrPalaThr 286  
Qy 243 ATTTTGGATTACAGGCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAA 302  
Db 287 IleLeuAspSerAlaGlyProLeuProAspTyThrProGluIlePheSerPheGlu 306  
Qy 303 AGTACTATCGGATTTACATTTGATGATGATGATGATGATGATGATGATGATGATGATG 362  
Db 307 SerThrGlyPheThrLeuTyGlyMetLeuTyLysProHisAspLeuGlnProGly 326  
Qy 363 AGAATATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422  
Db 327 LysLysTyProThrValLeuPheIleTyGlyGlyProGln----- 340  
Qy 423 GATCAGGTGGAAGGACTCCAATATCTAGTCTTCGATATGATTTCAATTGACTTAGATCGT 482  
Db 340 ----- 340  
Qy 483 GTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTTCCCTGATGGCATTAAATGACAGG 542

```
Db 340 ----- 340
QY 543 TCAGATATCTCAGGGTGTCTATGTGGGGCCAGCTACTCTGTGGATCTTCTATGAT 602
Db 341 -----ValAlaIleAlaGlyAlaProValThrLeuTyrPheTyrAsp 355
QY 603 ACAGGATACAGGAACGTTATATGGTGTACCCCTGACCAAGATGAACAGGCTATTACTTA 662
Db 356 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu 375
QY 663 GGATCTGTGCGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTCTCTTA 722
Db 376 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeu 395
QY 723 CATGGTTCTCGGATGAGATGTCCATTTTGCACATACACAGTATATTAAGTGTCTCTTA 782
Db 396 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 415
QY 783 GTGAGGCTCGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACACAGATTAAGA 842
Db 416 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 435
QY 843 GTTCCCTGAATCGGGAGACATATGAACTGCATCTTTTGCACACTTCAAGAAACCTT 902
Db 436 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 455
QY 903 GGAATCAGCTATTCTGCTCTTAAAGTGATA 932
Db 456 GlySerArgIleAlaLeuLysValIle 465

RESULT 6
US-09-976-674-3
; Sequence 3, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-3

Alignment Score:
Pred. No.: 1,04e-120 Length: 863
Score: 1093.50 Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1% Indels: 49
DB: 2 Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-3 (1-863)
QY 3 TTTCAGGACCAAGATCCCTTTAGAGCATCACTGTAGTGTAGTGTAGTAAAT 62
Db 515 PheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValIserTyrGluAla 534
QY 63 CQCTGAGAGGTGACAGGCTGACTGACCGGTGCTACTCATCTTCTGTCTGTCAGTCAG 122
Db 535 AlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGln 554
QY 123 CACTGTGACTCTTTTATAAGTAAAGTATAGTAAACAGAAATCCCACTGTGTCCCTT 182
::: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
555 AsnPheAspMetPheValSerHisTyrSerValSerValSerThrProCysValHisVal 574
183 TACAAGTATCAAGTCTGAAGATGACCCAACTGTGAAACAAAGAAATTTTGGCCACC 242
575 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTyrPalaser 594
243 ATTTTGGATTCACAGCGGTCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGA 302
595 MetMetGluAlaAlaSerCysProProAspTyrValProProGluIlePheHisPheHis 614
303 AGTACTACTGGATTTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 362
615 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 634
363 AAGAAATATCTACTGTGCTTTCATATATGTTGGTGTCTCAG----- 404
635 LysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 654
404 ----- 404
655 SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 674
404 ----- 404
675 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 694
405 -----GGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTACTTTCT 455
695 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 714
456 CGATATGATTTTCACTTGTAGATCGTGTGGGATCCACGGCTGTCTCTATGAGGATAC 515
715 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrPsrTyrGlyGlyPhe 734
516 CTCTCCCTGATGGCATTAATGACAGAGTCAAGATATCTTCAGGTTGTCTATTGCTGGGCC 575
735 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 754
576 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAAACGTTATATGTTGGTCC 635
755 ProValThrValTyrMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 774
636 GACCAGATGAACAGGCTATTACTTGAATCTGTGGCCATCAAGCAGAAAGTTCCCTCC 695
775 GluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 794
696 TCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGCTCCATTTTGA 755
795 AsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePhe 814
756 CATACCAATATATTACTAGTGTGTTTGTAGTGGGCTGGAAAGCCATATGATTTTACAGATC 815
815 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 834
816 TATCTCAGGAGACACACATAGAGTTCCTGAATCGGAGAACATTATGAACTGAT 875
835 TyrProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThr 854
876 CTTTGTGACTTCTTCAAGAAACCTT 902
855 LeuLeuHisPheLeuGlnGluTyrLeu 863

RESULT 7
US-09-976-674-23
; Sequence 23, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
```

```
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 23
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-23

Alignment Scores:
Pred. No.: 1,05e-120 Length: 892
Score: 1093.50 Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1% Indels: 49
DB: Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-23 (1-892)

QY 3 TTGAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAAT 62
DB 544 PheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGluAla 563
QY 63 CCTGGAGGTGACAGGCTGACCGCTGCTACTACATCTTCTGTCATCAGTCAG 122
DB 564 AlaGlyGluLeuValArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGln 583
QY 123 CACTGTGACTCTTTATAGTAAGTATAGTAAACAGAGAATCCACACTGTGTCCCTT 182
DB 584 AsnPheAspMetPheValSerHisTyrSerValSerThrProCysValHisVal 603
QY 183 TACAAGCTATCAAGTCTTGAAGATGACCCAACTTCGCAAAACAAAGAAATTTGGGCCACC 242
DB 604 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheThrPalaSer 623
QY 243 ATTTGGATTACGACAGGTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 302
DB 624 MetMetGluAlaAlaSerCysProProAspTyrValProProGluLeuPheHisPheHis 643
QY 303 AGTACTACTGGATTTACATTTGATGGATGCTTACAGGCTCATGATCTACAGCTGGA 362
DB 644 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 663
QY 363 AAGAAATATCTACTGTGCTGTTCATATATATGTTGCTCTCAG ----- 404
DB 664 LysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 683
QY 404 ----- 404
DB 684 SerPheLysGlyLysLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 703
QY 404 ----- 404
DB 704 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 723
QY 405 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAAATCTAGCTTCT 455
DB 724 AsnGlnMetGlyGlnValGluLeuGlnValGlnValGlnLeuGlnPheValAlaGlu 743
QY 456 CGATATGATTTCATTGACTTATAGATCGTGTGGGCATCCACGGCTGTGCTTATGAGGATAC 515
DB 744 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGlyPhe 763
QY 516 CTCCTCCTGATGGCTAATATGACAGAGGTGAGATATCTTCAGGGTGTGCTATTGTGGGGCC 575
DB 764 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 783
QY 576 CAGTCACTCTGTGATCTTCTATGATACAGATACAGGAACCTTATATGGGTCACTCT 635
DB 784 ProValThrValTyrMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 803

; Sequence 27, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 27
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-27

Alignment Scores:
Pred. No.: 1,05e-120 Length: 892
Score: 1093.50 Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1% Indels: 49
DB: Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-27 (1-892)

QY 3 TTGAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACGTAGTACGTAGTAAAT 62
DB 544 PheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGluAla 563
QY 63 CCTGGAGGTGACAGGCTGACCGTGGCTACTACATCTTCTGTCATCAGTCAG 122
DB 564 AlaGlyGluLeuValArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGln 583
QY 123 CACTGTGACTCTTTATAGTAAGTATAGTAAACAGAGAATCCACACTGTGTCCCTT 182
DB 584 AsnPheAspMetPheValSerHisTyrSerValSerThrProCysValHisVal 603
QY 183 TACAAGCTATCAAGTCTTGAAGATGACCCAACTTCGCAAAACAAAGAAATTTGGGCCACC 242
DB 604 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheThrPalaSer 623
QY 243 ATTTGGATTACGACAGGTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 302
DB 624 MetMetGluAlaAlaSerCysProProAspTyrValProProGluLeuPheHisPheHis 643
QY 303 AGTACTACTGGATTTACATTTGATGGATGCTTACAGGCTCATGATCTACAGCTGGA 362
DB 644 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 663
QY 363 AAGAAATATCTACTGTGCTGTTCATATATATGTTGCTCTCAG ----- 404
DB 664 LysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 683
QY 404 ----- 404
DB 684 SerPheLysGlyLysLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 703
QY 404 ----- 404
DB 704 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 723
QY 405 -----GGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAAATCTAGCTTCT 455
DB 724 AsnGlnMetGlyGlnValGluLeuGlnValGlnValGlnLeuGlnPheValAlaGlu 743
QY 456 CGATATGATTTCATTGACTTATAGATCGTGTGGGCATCCACGGCTGTGCTTATGAGGATAC 515
DB 744 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGlyPhe 763
QY 516 CTCCTCCTGATGGCTAATATGACAGAGGTGAGATATCTTCAGGGTGTGCTATTGTGGGGCC 575
DB 764 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 783
QY 576 CAGTCACTCTGTGATCTTCTATGATACAGATACAGGAACCTTATATGGGTCACTCT 635
DB 784 ProValThrValTyrMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 803
```

```
QY 303 AQTACTACTGGATTTTACATTGTATGGGATGCTCTACAAGCCTCATGATCTTACAGCCTGGA 362
   ::::
Db 644 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 663
   ::::
QY 363 AAGAATATCTTACTGCTGCTTTCATATATGGTGGTCTCAG----- 404
   ::::
Db 664 LysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 683
   ::::
QY 404 ----- 404
Db 684 SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 703
   ::::
QY 404 ----- 404
Db 704 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 723
   ::::
QY 405 -----GGTCAATAGAAATTCACGATCAGTGGAGGACTCCCAATATCTAGTCTTCT 455
   ::::
Db 724 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 743
   ::::
QY 456 CGATATGATTTTCATTGACTTATGATCGTGGGATCCACGGCTGGTCTTATGGAGATAC 515
   ::::
Db 744 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGlyPhe 763
   ::::
QY 516 CTCTCCCTGATGGCAATATGCAGAGGTTCAGATATCTTCAGGGTTCCTATTGCTGGGCC 575
   ::::
Db 764 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 783
   ::::
QY 576 CGAGTCACCTGTGTGATCTTCTATGATACAGGATACACGGAAGTTATATGGTTCACCT 635
   ::::
Db 784 ProValThrValTyrMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 803
   ::::
QY 636 GACCAAGATGAACAGCGCTATTACTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCC 695
   ::::
Db 804 GluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 823
   ::::
QY 696 TCTGAACAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTCCATTTTGCA 755
   ::::
Db 824 AsnGluProAsnArgLeuLeuIleuHisGlyPheLeuAspGluAsnValHisPhePhe 843
   ::::
QY 756 CATACAGTATATTTACTGAGTTTTTACTGAGGCTGGAAAGCCATATGATTTTACAGATC 815
   ::::
Db 844 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 863
   ::::
QY 816 TATCTCAGGACGACACAGCATAGAGTTCTCAATCGGAGAACATATGATGACATGCAT 875
   ::::
Db 864 TyrProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThr 883
   ::::
QY 876 CTTTTGCACCTACTCTCAAGAAACCTT 902
   ::::
Db 884 LeuLeuHisPheLeuGlnGluTyrLeu 892
   ::::
```

## RESULT 9

```
US-10-070-464-7
; Sequence 7, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1993-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7
```

```
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-7
```

## Alignment Scores:

Pred. No.:	4,31e-120	Length:	360
Score:	1086.50	Matches:	204
Percent Similarity:	80.6%	Conservative:	0
Best Local Similarity:	80.6%	Mismatches:	0
Query Match:	50.8%	Indels:	49
DB:	2	Gaps:	1

US-10-825-632-4 (1-1197) x US-10-070-464-7 (1-360)

```
QY 3 TTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCACCTGTAGTAGTCAGTTACGTAAAT 62
   ::::
Db 108 PheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSerTyrValAsn 127
   ::::
QY 63 CTGGAGAGGTGACAAAGGCTGACTGACCGTGGTACTACTCACATTTCTGCTGCATCAGTCAG 122
   ::::
Db 128 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysSileSerGln 147
   ::::
QY 123 CACTGTGACTTCTTTATAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGTCCTT 182
   ::::
Db 148 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 167
   ::::
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGGAATTTTGGGCCACC 242
   ::::
Db 168 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 187
   ::::
QY 243 ATTTTGGATTACAGAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
   ::::
Db 188 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 207
   ::::
QY 303 AGTACTACTGGATTACATTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGA 362
   ::::
Db 208 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 227
   ::::
QY 363 AAGAAATATCTCTACTGCTGTTCATATATGTTGCTCTCTCAG----- 404
   ::::
Db 228 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 247
   ::::
QY 404 ----- 404
Db 248 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 267
   ::::
QY 404 ----- 404
Db 268 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 287
   ::::
QY 405 -----GGTCAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGTCTCT 455
   ::::
Db 288 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 307
   ::::
QY 456 CGATATGATTTTCATTGACTTATGATCGTGGGATCCACGGCTGCTCTCTATGGAGGATAC 515
   ::::
Db 308 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 327
   ::::
QY 516 CTCTCCCTGATGGCAATATGAGAGGTCAGATATCTTACAGGTTGCTTATGCTGGGGCC 575
   ::::
Db 328 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 347
   ::::
QY 576 CCAGTCACTGTGGGATCTTCTATGATACAGGATACAG 614
   ::::
Db 348 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 360
   ::::
```

## RESULT 10

```
US-09-976-674-33
; Sequence 33, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
```

```
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 33
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-33

Alignment Scores:
Pred. No.: 7,14e-120 Length: 879
Score: 1086.50 Matches: 199
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.2% Mismatches: 58
Query Match: 50.8% Indels: 49
DB: 2 Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-33 (1-879)
QY 6 GAAGCACCAGAGTCCCTTTAGAGCATCCAGTGTACGTAGTACGTAAATCCT 65
DB 532 LysGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAlaAla 551
QY 66 GGAGAGGTGACCAAGCTCAGTCCGCTGCTACTCATTCTTGGTGCATCAGTACGAC 125
DB 552 GlyGluIleValArgLeuThrProGlyPheSerHisSerCysSerMetSerGlnAsn 571
QY 126 TGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACGTGTGCTTAC 185
DB 572 PheAspMetPheValSerHisTyrSerValSerThrProCysValHisValTyr 591
QY 186 AAGCTATCAAGTCTCTGAAGATGACCACTTGCACCAACAAAGAAATTTGGCCACCAT 245
DB 592 LysLeuSerGlyProAspAspAspProLeuHisGlnProArgPheTrpAlaSerMet 611
QY 246 TTGGATTACAGAGGTCTCTCTGCTACTACTCTCCAGAAATTTCTCTTTGAAAGT 305
DB 612 MetGluAlaAlaSerCysProAspTyrValProGluIlePheHisPheHisThr 631
QY 306 ACTACTGATTTACATGTTATGGATCTCTACAGCTCATGATCTACAGCTGGAAG 365
DB 632 ArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLys 651
QY 366 AAATATCTACTGCTGTTTCATATATGGTGGTCTCAG----- 404
DB 652 LysHisProThrValLeuPheValTyrGlyProGlnValGlnLeuValAsnSer 671
QY 404 ----- 404
DB 672 PheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValVal 691
QY 404 ----- 404
DB 692 ValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsn 711
QY 405 -----GGTCAATAGAAATGACGATCAGGTGGAAGGACTCAATATCTAGCTTCTCGA 458
DB 712 GlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLys 731
QY 459 TATGATTTCAATTGACTTAGATCGTGTGGGCATCCAGCGTGGTCTCTATGGAGGATACCTC 518
DB 732 TyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGlyPheLeu 751
QY 519 TCCCTGATGGCAATTAAGCAGAGGTACGATATCTTCAGGGTTCCTATTGCTGGGCCCA 578
DB 519 TCCCTGATGGCAATTAAGCAGAGGTACGATATCTTCAGGGTTCCTATTGCTGGGCCCA 578
```

```
DB 752 SerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAlaPro 771
QY 579 GTCACTCTGGGATCTTCTATGATACAGGATACAGGACCTTATATGGTGCACCCCTGAC 638
DB 772 ValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValProGlu 791
QY 639 CAGAAATGAACGGCTATTACTTAGACTGTGGCCATGCGAAGCAAGAAAGTTCCCTCT 698
DB 792 AsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuProAsn 811
QY 699 GAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGATGTCATTTTGCACAT 758
DB 812 GluProAsnArgLeuLeuIleHisGlyPheLeuAspGluAsnValHisPhePheHis 831
QY 759 ACCAGTATATTACTGAGTGTCTTACTAGGGCTGGAAGCCATATGATTTTACAGATCTAT 818
DB 832 ThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIleTyr 851
QY 819 CCTCAGGAGAGACACAGCATAAGAGTTCTGAATCGGAGAACATTTATGAAGTGCATCTT 878
DB 852 ProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThrLeu 871
QY 879 TTGCACTACCTTCAAGAAACCTT 902
DB 872 LeuHisPheLeuGlnGluTyrLeu 879

RESULT 11
US-09-976-674-35
; Sequence 35, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 35
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-35

Alignment Scores:
Pred. No.: 7,14e-120 Length: 879
Score: 1086.50 Matches: 199
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.2% Mismatches: 58
Query Match: 50.8% Indels: 49
DB: 2 Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-35 (1-879)
QY 6 GAAGCACCAGAGTCCCTTTAGAGCATCCAGTGTACGTAGTACGTAAATCCT 65
DB 532 LysGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAlaAla 551
QY 66 GGAGAGGTGACCAAGCTCAGTCCGCTGCTACTCATTCTTGGTGCATCAGTACGAC 125
DB 552 GlyGluIleValArgLeuThrProGlyPheSerHisSerCysSerMetSerGlnAsn 571
QY 126 TGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACGTGTGCTTAC 185
DB 572 PheAspMetPheValSerHisTyrSerValSerThrProCysValHisValTyr 591
QY 186 AAGCTATCAAGTCTCTGAAGATGACCACTTGCACCAACAAAGAAATTTGGCCACCAT 245
DB 592 LysLeuSerGlyProAspAspAspProLeuHisGlnProArgPheTrpAlaSerMet 611
QY 246 TTGGATTACAGAGGTCTCTCTGCTACTACTCTCCAGAAATTTCTCTTTGAAAGT 305
DB 612 MetGluAlaAlaSerCysProAspTyrValProGluIlePheHisPheHisThr 631
QY 306 ACTACTGATTTACATGTTATGGATCTCTACAGCTCATGATCTACAGCTGGAAG 365
DB 632 ArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLys 651
QY 366 AAATATCTACTGCTGTTTCATATATGGTGGTCTCAG----- 404
DB 652 LysHisProThrValLeuPheValTyrGlyProGlnValGlnLeuValAsnSer 671
QY 404 ----- 404
DB 672 PheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValVal 691
QY 404 ----- 404
DB 692 ValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsn 711
QY 405 -----GGTCAATAGAAATGACGATCAGGTGGAAGGACTCAATATCTAGCTTCTCGA 458
DB 712 GlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLys 731
QY 459 TATGATTTCAATTGACTTAGATCGTGTGGGCATCCAGCGTGGTCTCTATGGAGGATACCTC 518
DB 732 TyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGlyPheLeu 751
QY 519 TCCCTGATGGCAATTAAGCAGAGGTACGATATCTTCAGGGTTCCTATTGCTGGGCCCA 578
DB 519 TCCCTGATGGCAATTAAGCAGAGGTACGATATCTTCAGGGTTCCTATTGCTGGGCCCA 578
```







Db 544 PheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrrValValSerTyrrGluAla 563  
QY 63 CCTGAGAGGTGACAAAGCTGACCTGGTGGTACTCACATTCTTGTGTCATCAGTCAG 122  
Db 564 AlaGlyGluIleValThrProGlyPheSerHisSerCysSerMetSerGln 583  
QY 123 CACTGTGACTTCTTATAAGTAAATAGTATAGTACCAAGAAATCCACACTGTGTGCCCTT 182  
Db 584 AsnPheAspMetPheValSerHisTyrrSerSerValSerThrProCysValHisVal 603  
QY 183 TACAAGCTTATCAAGTCTCAAGATGACCCAACTTGCACAAACAAAGAAATTTGGCCACC 242  
Db 604 TyrrLysLeuSerGlyProAspAspAspProLeuHisLysGlnProArgPheIrrPalaSer 623  
QY 243 ATTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAAATTTCTCTTTTGA 302  
Db 624 MetMetGluAlaAlaSerCysProProAspTyrrValProGluIlePheHisPheHis 643  
QY 303 AGTACTACTGGATTTAGATTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362  
Db 644 ThrArgSerAspValArgLeuTyrrGlyMetIleTyrrLysProHisAlaLeuGlnProGly 663  
QY 363 AAGAAATATCTACTGCTGCTTCATATATGCTGCTCAG----- 404  
Db 664 LysLysHisProThrValLeuPheValTyrrGlyGlyProGlnValGlnLeuValAsnAsn 683  
QY 404 ----- 404  
Db 684 SerPheLysGlyIleLysTyrrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrrAlaVal 703  
QY 404 ----- 404  
Db 704 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 723  
QY 405 -----GTCAAATAGAAATTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCT 455  
Db 724 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 743  
QY 456 CDAATATGATTTCATTGACTTAGATCGTGTGGGCATCCAGCGCTGCTTATGGAGGATAC 515  
Db 744 LysTyrrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrrSerTyrrGlyPhe 763  
QY 516 CTCTCCCTGATGGCAATTAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575  
Db 764 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysAlaGlnProLeuAlaTyrr 783  
QY 576 CAGTCACCTCG 587  
Db 784 ProProArgLeu 787

Search completed: May 2, 2006, 01:42:08  
Job time : 53.6769 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	1680	78.5	310	3	US-09-593-955-4	Sequence 4, Appli
2	1680	78.5	310	4	US-10-825-632-3	Sequence 3, Appli
3	1645.5	76.9	882	3	US-09-976-674-1	Sequence 1, Appli
4	1645.5	76.9	882	4	US-10-054-776-2	Sequence 2, Appli
5	1645.5	76.9	882	4	US-10-170-789-38	Sequence 38, Appli
6	1645.5	76.9	882	4	US-10-311-035-9	Sequence 9, Appli
7	1645.5	76.9	882	4	US-10-072-013-622	Sequence 622, App
8	1645.5	76.9	882	4	US-10-415-123-6	Sequence 6, Appli
9	1645.5	76.9	882	4	US-10-825-632-1	Sequence 1, Appli
10	1645.5	76.9	882	5	US-10-982-513-1	Sequence 1, Appli
11	1605.5	75.0	883	4	US-10-072-013-621	Sequence 621, App

Qy	3	TTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTGACGTAGTCAGTTACGTTAAAT	62
Db	1	PheGluGlyThrIysaspSerProLeuGluHisIseuTyValValSerTyValAsn	20
Qy	63	CCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTATCTCACATTCCTTGCTGCATCAGTCAG	12

```
Db 21 ProGlyGluValThrArgLeuThrAspArgGlyThrSerHisSerCysHisSerGln 40
Qy 123 CACTGTGACTCTTTTATAAGTATAGTAACACAGAGAAATCCACACTGTGTGCCCTT 182
Db 41 HisCysAspPhePheIleSerLysThrSerAsnGlnLysAsnProHisCysValSerLeu 60
Qy 183 TACAAGCTATCAAGTCTCAAGATGACCAACTTGCAAAACAAAGGAATTTTGGGCACC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTpaLaThr 80
Qy 243 ATTTTGGATTACAGCAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspThrProProGluIlePheSerPheGlu 100
Qy 303 AGTACTACTGGATTTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCTGGA 362
Db 101 SerThrThrGlyPheThrLeuThrGlyMetLeuThrLysProHisAspLeuGlnProGly 120
Qy 363 AAGAAATATCCTACTGTCTGCTCATATATATGTTGCTCTCAGGCTCAAAATAGAAATGAC 422
Db 121 LysLysThrProThrValLeuPheIleThrGlyGlyProGlnGlyGlnIleGluIleAsp 140
Qy 423 GATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGATCTCT 482
Db 141 AspGlnValGluGlyLeuGlnThrLysLeuAlaSerArgThrAspPheIleAspLeuAspArg 160
Qy 483 GTGGGATCCACGGCTGTCTTATGAGAGATACCTCTCCCTGATGGATTAATGACAGG 542
Db 161 ValGlyIleHisGlySerThrGlyThrLysLeuSerLeuMetAlaLeuMetGlnArg 180
Qy 543 TCAGATATCTTCAAGGTTTCTATGCTGGGGCCCGACACCTCTGTGGATCTTCTATGAT 602
Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuThrPheThrAsp 200
Qy 603 ACAGGATACACGGAAGCTTTATATGGGTACCCCTGACCAAGAAATGACAGGCTATTCTTA 662
Db 201 ThrGlyThrThrGluArgThrMetGlyHisProAspGlnAsnGluGlnGlyThrThrLeu 220
Qy 663 GGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTTTA 722
Db 221 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 240
Qy 723 CATGGTTTCTGGATCAGAAATGTCATTTTGCACATACCATATATTAATTAATTAATTA 782
Db 241 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
Qy 783 GTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAGA 842
Db 261 ValArgAlaGlyLysProThrAspLeuGlnIleThrProGlnGluArgHisSerIleArg 280
Qy 843 GTTCTCGAATCGGAGAACATTAATGACCTGCACTTTTGGCACTACCTTCAAGAAACCTT 902
Db 281 ValProGluSerGlyGluHisThrGluLeuHisLeuLeuHisThrLeuGlnGluAsnLeu 300
Qy 903 GGATCAGTATTCCTCTCTTAAGTGATA 932
Db 301 GlySerArgIleAlaAlaLeuLysValIle 310
```

## RESULT 2

```
US-10-825-632-3
; Sequence 3, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCBSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
```

```
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-825-632-3
```

## Alignment Scores:

```
Pred. No.: 4,85e-169 Length: 310
Score: 1680.00 Matches: 310
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 78.5% Indels: 0
DB: Gaps: 0
```

US-10-825-632-4 (1-1197) x US-10-825-632-3 (1-310)

```
Qy 3 TTTGAGGCGACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTACGTAAAT 62
Db 1 PheGluGlyThrLysAspSerProLeuGluHisLeuThrValValSerThrValAsn 20
Qy 63 CTTGGAGAGGTGACRAGGCTGACTGACCGTGGCTACTCACATTTCTTCTGCATCAGTCAG 122
Db 21 ProGlyValThrArgLeuThrAspArgGlyThrSerHisSerCysValSerGln 40
Qy 123 CACTGTGACTCTTTTATAAGTATAGTAACACAGAGAAATCCACACTGTGTGCCCTT 182
Db 41 HisCysAspPhePheIleSerLysThrSerAsnGlnLysAsnProHisCysValSerLeu 60
Qy 183 TACAAGCTATCAAGTCTCTGAAAGATGACCCCACTTCCAAAACAAAGAAATTTTGGGCACC 242
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheThrPalaThr 80
Qy 243 ATTTTGGATTACAGCAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 81 IleLeuAspSerAlaGlyProLeuProAspThrProProGluIlePheSerPheGlu 100
Qy 303 AGTACTACTGGATTTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCTGGA 362
Db 101 SerThrThrGlyPheThrLeuThrGlyMetLeuThrLysProHisAspLeuGlnProGly 120
Qy 363 AAGAAATATCCTACTGTCTGCTCATATATGTTGCTCTCAGGCTCAAAATAGAAATGAC 422
Db 121 LysLysThrProThrValLeuPheIleThrGlyGlyProGlnGlyGlnIleGluIleAsp 140
Qy 423 GATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTATGATCGT 482
Db 141 AspGlnValGluGlyLeuGlnThrLysLeuAlaSerArgThrAspPheIleAspLeuAspArg 160
Qy 483 GTGGGATCCACGGCTGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGG 542
Db 161 ValGlyIleHisGlyThrSerThrGlyGlyThrLysLeuSerLeuMetAlaLeuMetGlnArg 180
Qy 543 TCAGATATCTTCAAGGTTTCTATGCTGGGTACCTGACCGAGATGACAGGCTTCTTATGAT 602
Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuThrPheThrAsp 200
Qy 603 ACAGGATACACGGAAGCTTTATGCTGCTACCTGACCGAATGACAGGCTTACTACTTA 662
Db 201 ThrGlyThrThrGluArgThrMetGlyHisProAspGlnAsnGluGlnGlyThrThrLeu 220
Qy 663 GGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTTTA 722
Db 221 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 240
Qy 723 CATGGTTTCTGGATCAGAAATGTCATTTTGCACATACCATATTAATTAATTAATTA 782
```

```
Db 241 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260
QY 783 GTGAGGCGTGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATAGA 842
Db 261 ValArgAlaGlyLysProTyrAspLeuGluInIleTyrProGlnGluArgHisSerIleArg 280
QY 843 GTTCCTGAATCGGAGAACATTATGAACTGCACTCTTTTGCACTACCTTCAAGAAAACCTT 902
Db 281 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 300
QY 903 GATCAGCTATTGCTGCTCTTAAACATGATA 932
Db 301 GlySerArgIleAlaAlaLeuLysValIle 310

RESULT 3
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1.

Alignment Scores:
Pred. No.: 3,21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservatives: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 3 Gaps: 1

US-10-825-632-4 (1-1197) x US-09-976-674-1 (1-882)
QY 3 TTGAAGCCACCAAGACTCCCTTTAGAGCATCACTGTAGTACGTAGTCAGTTACGTAAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 543
QY 63 CTGGAGAGGTGACAAAGCTGACGTGACCGTGGCTACTACATCTTCTGTCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 123 CACTGTGACTTCTTTAAGTAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGAAATTTGGGCCAAC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTGGATTACAGAGTCTCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGATTTACATTTGATGATGCTCTCAAGCCTCATGATCTACAGCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTTCAG----- 404
```

```
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GTTCAAAATAGAAATTTGACGATCAGGTGAGGAGGACTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CGATATGATTTTCATTTGACTTTAGATCGTGTGGGCATCCACGGCTGCTCTTATGGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
QY 516 CTCTCCCTGATGCGATTAAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CCAGTCACTCTGTGGATCTCTATGATACAGGATACACGGAACGTTATATGGTCACCCCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAGAAATGAACAGGGCTATTACTTATGAGTCTGTGGCCATCAACGACGAGAAAGTTCCCC 695
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACCACTATATCTAGCTTTTCTAGTGGGCTGGAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCTCTCAGGAGACACACAGATAGAGTTCTCTGAATCGGAGAACATTTATGAACATGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTGTGACACTCTTCAAGAAACCTTGGATCAGTATTCGCTGCTCTTAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuLeuGlySerArgIleAlaLeuLysValIle 882

RESULT 4
US-10-054-776-2
; Sequence 2, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-776-2

Alignment Scores:
Pred. No.: 3,21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservatives: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 4 Gaps: 1
```

```
US-10-825-632-4 (1-1197) x US-10-054-776-2 (1-882)
Qy 3 TTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAen 543
Qy 63 CTGGAGAGGTGACAAAGGTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
Qy 123 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCTTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSerLeu 583
Qy 183 TACAAGCTATCAAGTCTGAAGTACACCACTTGCAGAAATTTCTGGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
Qy 243 ATTTTGGATTGACAGGTCCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluLullePheSerPheGlu 623
Qy 303 AGTACTACTGGATTACATTGATGGGATGCTCTACAGCTCATGATCTACAGCTTGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 363 AAGAAATATCTACTGTGCTTCATATATATGTTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAenAen 663
Qy 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGlyTyrValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAenArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GCTCAAAATAGAAATTCACGATCAGGTGGAGGACTCCCAATATCTAGTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAepGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 456 CGATATGATTTCATTCATTAGATCGTGTGGCATCCACGGCTCGTCTATGAGAGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAaspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
Qy 516 CTCTCCCTGATGCATTAATGACAGGTCAGATATCTTCAGGGTGTGCTATTGCTGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 576 CCAGTCACTCTGTGGATCTCTATGATACAGGATACAGGAACGTTATATGSGTCCACCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
Qy 636 GACCAGAAATGACAGGGCTATTACTTAGGATCTGGCCATGCAAGCAGAAAGTTCCCC 695
Db 784 AspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
Qy 696 TCTGAACCAAAATCGTTTACTGCTTTACATGTTTCTGGATGAGAAATGTCCTATTGCA 755
Db 804 SerGluProAenArgLeuLeuLeuHisGlyPheLeuAspGluAenValHisPheAla 823
Qy 756 CATACCAAGTATATTACTAGTGTTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
Qy 816 TATCTCAGGAGACACAGCATCAAGATTCCTGAAATCGGGAGAACATTTATGAACATCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
Qy 876 CTTTGTGACTACCTTCAAGAAACCTTGGATCACCTATTGCTGTCTTAAAGTATGATA 932
Db 864 LeuLeuHisTyrLeuGlnAenLeuGlySerArgIleAlaAlaLeuLysValIle 882
```

---

```
RESULT 5
US-10-170-789-38
; Sequence 38, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 882
```



TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-170-789-38

Alignment Scores:

Pred. No.: 3, 21e-165 Length: 882  
Score: 1645.50 Matches: 310  
Percent Similarity: 86.4% Conservative: 0  
Best Local Similarity: 86.4% Mismatches: 0  
Query Match: 76.9% Indels: 49  
DB: 4 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-170-789-38 (1-882)

QY 3 TTGAAGCACAAGACTCCCTTTAGACATCACTGTACGTAGTCAGTTACGTAAT 62  
DB 524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyValValSerTyValAsn 543  
QY 63 CTGGAGAGGTGACAAGGCTGACTGACCGTGGTACTACATCTTGTGTCATGTCAG 122  
DB 544 ProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCysCysLeSerGln 563  
QY 123 CACTGTGACTCTTTATAAGTAAGTATAGTAACACAGAAGAACCCACACTGTGTGCCCTT 182  
DB 564 HisCysAspPhePheLeSerLysTySerAsnGlnLysAsnProHisCysValSerLeu 583  
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGAAATTTTGGGCCACC 242  
DB 584 TyLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
QY 243 ATTTGGATTGACAGAGTCTCTCTGACTATCTCTCCAGAAATTTTCTTTTGAA 302  
DB 604 IleLeuAspSerAlaGlyProLeuProAspTyThrProProGluIlePheSerPheGlu 623  
QY 303 AGTACTACTGATTACATTTGATGATGATGCTCTCAAGCCTCATGATCTACAGCTGGA 362  
DB 624 SerThrThrGlyPheThrLeuTyrglyMetLeuTyLysPheHisAspLeuGlnProGly 643  
QY 363 AAGAAATCTCTACTGTGCTTTCATATATGTTGGTCTCTCAG----- 404  
DB 644 LysLysTyProThrValLeuPheIleTyGlyGlyProGlnValGlnLeuValAsnAsn 663  
QY 404 ----- 404  
DB 664 ArgPheLysGlyValLysTyPheArgLeuAsnThrLeuAlaSerLeuGlyTyValVal 683  
QY 404 ----- 404  
DB 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
QY 405 -----GGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 455  
DB 704 TyLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyLeuAlaSer 723  
QY 456 CQATATGATTCATTGACTTAGATCGTGTGGGCATCCAGCGTGTCTCTATGGAGATAC 515  
DB 724 ArgTyAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyGlyGlyTy 743  
QY 516 CTCTCCCTGATGCAATTAATGACAGGTGAGATATCTTCAGGGTGTCTATTGCTGGGCC 575  
DB 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
QY 576 CQAGTCACTGTGATCTTCTATGATACAGGATACACGGAAGCTTATATGGTCACCCCT 635  
DB 764 ProValThrLeuTrpIlePheTyAspThrGlyTyThrGluArgTyMetGlyHisPro 783  
QY 636 GACCAAGATGACAGGGCTATTACTAGGATCTGTGGCCATCCAGCAGAGAAAGTCCCC 695  
DB 784 AspGlnAsnGluGlnGlyTyTyLeuGlySerValAlaMetGlnAlaGluLysPhePro 803  
QY 696 TCTGACCAAAATCGTTTACTGCTCTTACATGTTTCTGGATGAGAATGTTCATTTGCA 755  
DB 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823

RESULT 6

US-10-311-035-9  
; Sequence 9, Application US/10311035  
; Publication No. US20040023243A1

GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: YUE, Henry  
; APPLICANT: ELLIOTT, Vicki  
; APPLICANT: GANDHI, Ameena R.  
; APPLICANT: LAL, Preeti  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: DELEGEANE, Angelo M.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: NGUYEN, Dannel B.  
; APPLICANT: LEE, Ernestine A.  
; APPLICANT: HAFALTA, April  
; APPLICANT: KHAN, Farrah A.  
; APPLICANT: CHAWLA, Narinder K.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: LU, Dyung Aina M.  
; APPLICANT: ARVIZU, Chandra S.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: WALSH, Roderick T.  
; APPLICANT: AZIMZAI, Valda  
; APPLICANT: LU, Yan  
; APPLICANT: RAMKUMAR, Jayalaximi  
; APPLICANT: XU, Yuming  
; APPLICANT: REDDY, Roopa  
; APPLICANT: DAS, Depopriya  
; APPLICANT: KEARNEY, Liam  
; APPLICANT: KALLICK, Deborah A.  
; TITLE OF INVENTION: Proteases  
; FILE REFERENCE: PI-0123 PCT  
; CURRENT APPLICATION NUMBER: US/10/311,035  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946  
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1  
US-10-311-035-9

Alignment Scores:

Pred. No.: 3, 21e-165 Length: 882  
Score: 1645.50 Matches: 310  
Percent Similarity: 86.4% Conservative: 0  
Best Local Similarity: 86.4% Mismatches: 0  
Query Match: 76.9% Indels: 49  
DB: 4 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-311-035-9 (1-882)

QY 3 TTGAAGCACAAGACTCCCTTTAGACATCACTGTAGTCAGTTACGTAAT 62

Db	524	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAen	543	
Qy	63	CCTGGAGAGGTACAAAGGCTGACTGACCGTGGCTACTCACATTCTTGGCTGCATCAGTCAG	122	
Db	544	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysSileSerGln	563	
Qy	123	CACGTGACTCTTTTATAGTAAAGTATAGTAACAGAGAAGTAATCCACACTGTGTGTCCTTT	182	
Db	564	HisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSerLeu	583	
Qy	183	TACAAGCTATCAAGTCTCAAGATGACCAACTTGGCAAAACAAGGAATTTTGGGCCACC	242	
Db	584	TyrlsLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr	603	
Qy	243	ATTTGGATTACAGAGCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA	302	
Db	604	IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu	623	
Qy	303	AGTACTACTGGATTACATGTATGGGATGCTCTACAGCTCTATGATCTACAGCTGGA	362	
Db	624	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	643	
Qy	363	AGAAATATCTACTGCTGCTTCATATATATATGCTGCTCTAG	404	
Db	644	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAenAen	663	
Qy	404	-----	404	
Db	664	ArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGlyTyrValVal	683	
Qy	404	-----	404	
Db	684	ValValIleAspAenArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	703	
Qy	405	-----CGTCAAATAGAAATTCAGATCAGGTGGAGGACTCCAATATCTAGTTCT	455	
Db	704	TyrlsMetGlyGlnIleGluIleAspAepGlnValGluGlyLeuGlnTyrLeuAlaSer	723	
Qy	456	CGATGATGTTTCACTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGAGATAC	515	
Db	724	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyrGlyGlyTyr	743	
Qy	516	CTCTCCCTGATGCATTAAATGACAGAGGTACAGATATCTTCAGGGTTGCTATTGCTGGGCC	575	
Db	744	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763	
Qy	576	CCAGTCACTCTGTGGATCTCTATGATACAGGATACAGAACTTATATGGTCCACCT	635	
Db	764	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro	793	
Qy	636	GACCAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCCAAGCAAGAAAGTTCC	695	
Db	784	AspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro	803	
Qy	696	TCTGAACCAATCGTTTACTGCTCTTACATGTTTCTGGATGAGATGTCATTTTGGCA	755	
Db	804	SerGluProAenArgLeuLeuLeuHisGlyPheLeuAspGluAenValHisPheAla	823	
Qy	756	CATACAGTATATTACTAGTCTTTTATGAGGGCTGGAAAGCCCATATGATTTACAGATC	815	
Db	824	HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle	843	
Qy	816	TATCCTCAGGAGACACACATAGAGTTCCTGTAATCGGGAGAACATTTAAGACTCCAT	875	
Db	844	TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis	863	
Qy	876	CTTTTGGCTACCTTCAAGAAAACCTTGGATCACGTTATGCTGCTCTCTAAAAGTGATA	932	
Db	864	LeuLeuHisTyrLeuGlnGlnAenLeuLeuGlySerArgIleAlaAlaLeuLysValIle	882	
RESULT 7				
US-10-072-012-622				

Sequence 622, Application US/10072012

Publication No. US2004003493A1

GENERAL INFORMATION:

APPLICANT: Tchernev, Velizar

APPLICANT: Spytek, Kimberly

APPLICANT: Zerhusen, Bryan

APPLICANT: Patturajan, Meera

APPLICANT: Shimketa, Richard

APPLICANT: Li, Li

APPLICANT: Gangolli, Esha

APPLICANT: Padigar, Muralidhara

APPLICANT: Anderson, David W.

APPLICANT: Rastelli, Luca

APPLICANT: Miller, Charles E.

APPLICANT: Gerlach, Valerie

APPLICANT: Taupier Jr, Raymond J.

APPLICANT: Gusev, Vladimir Y.

APPLICANT: Colman, Steven D.

APPLICANT: Wolenc, Adam R.

APPLICANT: Pena, Carol E. A.

APPLICANT: Furtak, Katarzyna

APPLICANT: Grosse, William M.

APPLICANT: Alsebrook II, John P.

APPLICANT: Lepley, Denise M.

APPLICANT: Rieger, Daniel K.

APPLICANT: Burgess, Catherine E.

APPLICANT: Title of Invention: proteins and Nucleic Acids Encoding Same

FILE REFERENCE: 21402-958

CURRENT APPLICATION NUMBER: US/10/072,012

PRIOR FILING DATE: 2002-01-31

PRIOR APPLICATION NUMBER: 60/265,102

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 60/265,514

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,517

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,412

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,395

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/266,406

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 60/266,767

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: 60/267,057

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/266,975

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/267,459

PRIOR FILING DATE: 2001-02-08

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1391

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO: 622

LENGTH: 882

TYPE: PRT

ORGANISM: Homo sapiens

US-10-072-012-622

Alignment Scores:

Pred. No.: 3 21e-165

Score: 1645.50

Percent Similarity: 86.4%

Best Local Similarity: 86.4%

Query Match: 76.9%

DB: 4

Length: 882

Matches: 310

Conservative: 0

Mismatches: 49

Indels: 1

Gaps: 1

US-10-825-632-4 (1-1197) x US-10-072-012-622 (1-882)

Qy 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGTACGTACGTAGTAAAT 62

Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAen 543

```
Qy 63 CTTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAATCTTCTGCTGATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 563
Qy 123 CACTGTGACTCTTTTATAAGTAGTAGTAACACAGAAGATCCACACTGTGTGTCCTTT 182
Db 564 HisCysAspPhePheLeuSerTyrSerAsnGlnLysAsnProHisCysValSerLeu 593
Qy 183 TACAAGCTATCAAGTCTCAAGATGACCCCACTTGCACAAACAAAGAAATTTTGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrAlaThr 603
Qy 243 ATTTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPheSerPheGlu 623
Qy 303 AGTACTACTGGATTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 363 AAGAAATATCTACTGCTGCTGTTCATATATATGCTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GGTCAAAATAGAAATTCAGATCGAGTGGAGGACTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 456 CQATATGATTTCAATGATTCAGTCTAGATCCTGTGGCCTCCAGGCTGCTCTATGGAGTAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPsrTyrGlyTyr 743
Qy 876 CTTTGGCACTTACCTCAAGAAACCTTGGATCAGCTATGCTGCTCTTAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIleAlaLeuLeuLysValIle 882
```

RESULT 8

US-10-415-122-6  
; Sequence 6, Application US/10415122  
; Publication No. US20040053369A1  
; GENERAL INFORMATION:

```
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASHS
; FILE REFERENCE: FP15217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-122-6

Alignment Scores:
Pred. No.: 3,21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 4 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-415-122-6 (1-882)
Qy 3 TTTGAAGCGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy 63 CTTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAATCTTCTGCTGATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 563
Qy 123 CACTGTGACTCTTTTATAAGTAGTAGTAACACAGAAGATCCACACTGTGTGTCCTTT 182
Db 564 HisCysAspPhePheLeuSerTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 183 TACAAGCTATCAAGTCTCTGAGATGACCACTTGCACAAACAAAGAAATTTTGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTyrAlaThr 603
Qy 243 ATTTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPheSerPheGlu 623
Qy 303 AGTACTACTGGATTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 363 AAGAAATATCTACTGCTGCTGTTCATATATGCTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GGTCAAAATAGAAATTCAGATCGAGTGGAGGACTCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 456 CQATATGATTTCAATGATTCAGTCTAGATCCTGTGGCCTCCAGGCTGCTCTATGGAGTAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPsrTyrGlyTyr 743
Qy 516 CTTTGGCACTTACCTCAAGAAACCTTGGATCAGCTATGCTGCTCTTAAAGTGATA 575
Db 744 LeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIleAlaLeuLeuLysValIle 763
Qy 576 CCACTCTACTCTGTGGATCTTCTTATGATACAGGATACAGGAACTGTTATGCGTCACT 635
```

Db 764 PrbValThrLeuTrpIlePheTyrAspThrThrGlyTyrThrGluArgTyrMetGlyHisPro 783  
QY 636 GACCAGATGACAGGGCTATTACTTAGGATCTGTGGCCATGACAGCAGAAAGTTCCCC 695  
Db 784 ApGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803  
QY 696 TCTGAACCAAAATCGTTTACTGCTTACATCGTTTCTGAGAGAAATGTCATTTGCA 755  
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823  
QY 756 CATACAGPATATTACTAGTTTATAGTGGGCTCGAAGCCATATGATTTACAGATC 815  
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843  
QY 816 TATCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTATGAACCTGCAT 875  
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863  
QY 876 CTTTTCGCACTACTCTCAAGAAACCTTGGATCAGTATTGCTCTCTAAAGTGATA 932  
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882

## RESULT 9

US-10-825-632-1  
; Sequence 1, Application US/10825632  
; Publication No. US20040191826A1  
; GENERAL INFORMATION:  
; APPLICANT: ABSORT, Catherine Anne  
; APPLICANT: CORRELL, Mark Douglas  
; TITLE OF INVENTION: DIFEPTIDYL PEPTIDASES  
; FILE REFERENCE: FCSB-100-Div. 1  
; CURRENT APPLICATION NUMBER: US/10/825.632  
; CURRENT FILING DATE: 2004-04-15  
; PRIOR APPLICATION NUMBER: US 10/070,464  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-825-632-1

Alignment Scores:  
Pred. No.: 3.21e-165 Length: 882  
Score: 1645.50 Matches: 310  
Percent Similarity: 86.4% Conservative: 0  
Best Local Similarity: 86.4% Mismatches: 49  
Query Match: 76.9% Indels: 49  
DB: 4 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-825-632-1 (1-882)

QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTACTGTTACGTAAAT 62  
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543  
QY 63 CTTGAGAGGTGACAGGCTCACTGACCGTGGCTACTCACATTTCTTCTGCTCACTGCTAG 122  
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysLysSerGln 563  
QY 123 CACTGTGACTTCTTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTT 182  
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
QY 183 TACAAGCTATCAAGTCTCAAGATGACCACTTGCACCAAGAAATTTGGGGCCACC 242

Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
QY 243 ATTTTCGATTCAGAGCTCTCTCTCTGACTATACTCTCCAGAAATTTCTCTTTTGAA 302  
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
QY 303 AGTACTACTGGATTTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTGGA 362  
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
QY 363 AAGAAATATCTACTGTCTGTATATATATATGTTGCTCTCAG----- 404  
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663  
QY 404 ----- 404  
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
QY 404 ----- 404  
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
QY 405 -----GGTCAAATAGAAATTCACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 455  
Db 704 TyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
QY 456 CGATATGATTTCATTGACTTAGATCTGTGGGCATCCAGCGCTGCTCTATGGAGATAC 515  
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyrGlyTyr 743  
QY 516 CTCTCCCTGATGGCATTAATGAGAGGTTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575  
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
QY 576 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTTATATGGGTCACCT 635  
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783  
QY 636 GACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCATTCGACAGCAGAAAGTTCCCC 695  
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803  
QY 696 TCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGAAATGTCATTTTGCA 755  
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823  
QY 756 CATACAGTATATTACTAGTTTCTTGTAGGCTGGAAGCCATATGATTTACAGATC 815  
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843  
QY 816 TATCTCAGGAGACACAGCATAGACTTCTCTGAATCGGAGAACATTATGAACCTGCAT 875  
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863  
QY 876 CTTTTCGCACTACTCTCAAGAAACCTTGGATCAGTATTGCTCTCTAAAGTGATA 932  
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882

## RESULT 10

US-10-982-512-1  
; Sequence 1, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982.512  
; CURRENT FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12

```
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-1

Alignment Scores:
Pred. No.: 3,21e-165 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 5 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-982-512-1 (1-882)
QY 3 TTTGAAGCCACCAAGACTCCCTTTAGAGCATCACTGTAGTAGTCAGTTACGTAAAT 62
DB 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 63 CTGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCTTGTGTCATCAGTCAG 122
DB 544 ProGluGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 123 CACTGTGACTCTTTATAGTATAGTATACACAGAGAATCCACACTGTGTCTCCCTT 182
DB 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAGCTATCAAGTCTGAGATGACCCCACTTCACAAACAAAGAAATTTTGGCCACC 242
DB 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTGGATTACAGAGTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGA 302
DB 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
QY 303 AdTACTACTGGATTTACATTGTATGGATGCTCTCAAGCCCTCATGATCTACAGCCCTGGA 362
DB 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGTGCTTTCATATATATGTTGCTCTCAG----- 404
DB 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
DB 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
DB 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GGTCAAAATAGAAATTTGACGATCAGGTGAGGAGGACTCCAATATCTAGCTTCT 455
DB 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CdATATGATTTCACTGACTTAGATCGTGTGGGCATCCAGCGCTGCTCTCTATGGAGGATAC 515
DB 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743
QY 516 CTTCTCCCTGATGCAATTAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
DB 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CCAGTCACTCTGTGGATCTCTTATATGATACAGATACAGGAACGTTATATGGTCAACCT 635
DB 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAGATGACAGGGCTATTACTTAGGATCTGTGGCCATCGACAGCAAGAAAGTTCCCC 695
```

```
784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY TCTGAACCAAAATCGTTTACTGCTTACATGCTTCTCGGATGAGAAATGTCATTGCA 755
DB SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACCAAGTATATTACTGTAGTTTTTTAGTGAGGCTGGAAAGCCCATATGATTACAGATC 815
DB HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCTCAGGAGAGACACAGCATAGAGTTCTGTAATCGGAGAACATTATGAATGCAAT 875
DB TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTGTGACTACCTTCAAGAAAACCTTGGATCACGATTTGCTGCTCTAAAAGTGTATA 932
DB LeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIleAlaLeuLysValIle 882

RESULT 11
US-10-072-012-621
; Sequence 621, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Baha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
```



```
QY 303 AGTACTACTGGAATTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 307 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 326
QY 363 AAGAAATACCTACTGCTGCTCATATATGCTGCTCCTCAGGCTCAATAGAAATTGAC 422
Db 327 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGln----- 340
QY 423 GATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTTAGATCGT 482
Db 340 ----- 340
QY 483 GTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCTCATGGCATTTAATGACAGG 542
Db 340 ----- 340
QY 543 TCAGATATCTTCAGGGTGTCTATTGCTGGGGCCCGCAGTCACCTCTGCTGATCTTCTATGAT 602
Db 341 -----ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 355
QY 603 ACAGGATACACGGAACGTTATATGGGTCACTGACAGCAATGAACAGGCTATTACTTTA 662
Db 356 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu 375
QY 663 GdATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATFCGTTTACTCTCTTA 722
Db 376 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 395
QY 723 CATGGTTTCCTGGATGAGATGTCCTATTTTGCACATACCATGATATATTACTAGTTTAA 782
Db 396 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 415
QY 783 GTGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGACACAGCATAGA 842
Db 416 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 435
QY 843 GTTCTGAATCGGAGAACATTATGAACCTGCATCTTTTGCACATCTTCAAGAAAACTT 902
Db 436 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 455
QY 903 GGATCAGCTATTGCTGCTCTAAAAGTGATA 932
Db 456 GlySerArgIleAlaLeuLysValIle 465
```

RESULT 13

US-10-275-505-2  
; Sequence 2, Application US/10275505  
; Publication No. US20040081961A1  
; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: DELEGEANE, Angelo M.; LAL, Preeti G.  
; APPLICANT: HAFALIA, April J.A.; PATTERSON, Chandra  
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam  
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.  
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.  
; APPLICANT: AZIMZAI, Valda; ELLIOTT, Vicki S.  
; APPLICANT: NGUYEN, Dannel B.; GANDHI, Ameena R.  
; APPLICANT: YANG, Junning; HERNANDEZ, Roberto  
; APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.  
; APPLICANT: REDDY, Roopa M.; YUE, Henry  
; APPLICANT: TANG, Y. Tom

; TITLE OF INVENTION: PROTEASES

; FILE REFERENCE: PI-0085 USN

; CURRENT APPLICATION NUMBER: US/10/275, 505

; CURRENT FILING DATE: 2002-11-04

; PRIOR APPLICATION NUMBER: PCT/US01/14651

; PRIOR FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: 60/209, 402

; PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: 60/207, 477

; PRIOR FILING DATE: 2000-05-25

; PRIOR APPLICATION NUMBER: 60/205, 803

; PRIOR FILING DATE: 2000-05-17

; PRIOR APPLICATION NUMBER: 60/203, 566  
; PRIOR FILING DATE: 2000-05-11  
; PRIOR APPLICATION NUMBER: 60/202, 082  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ. ID NOS: 28  
; SOFTWARE: PERL Program  
; SEQ ID NO 2  
; LENGTH: 580  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 376067CDI  
US-10-275-505-2

Alignment Scores:

Pred. No.: 3, 22e-136 Length: 580  
Score: 1372.00 Matches: 266  
Percent Similarity: 87.1% Conservative: 11  
Best Local Similarity: 83.6% Mismatches: 19  
Query Match: 64.1% Indels: 22  
DB: 4 Gaps: 6

US-10-825-632-4 (1-1197) x US-10-275-505-2 (1-580)

```
QY 3 TTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTACGTTAGCTAAAT 62
Db 277 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 296
QY 63 CCTGGAGAGTGACAGGCTGACTGACCGGTGGCTACTACATCTTGTGTCGATCAGTCAG 122
Db 297 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 316
QY 123 CACTGTGACATCTTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 182
Db 317 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 336
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCAGAAACAAAGAAATTTTGGGCCACC 242
Db 337 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 356
QY 243 ATTTTGAATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 357 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 376
QY 303 AGTACTACTCGATTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCTGGA 362
Db 377 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 396
QY 363 AAGAAATATCTACTGTGCTGTTTCATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 422
Db 397 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu---ValAsn 415
QY 423 GATCAGGTGGAAGACTCCCAATAT-----CTAGCTTCT---CGATATGAT 464
Db 416 AsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal 435
QY 465 TTCATTGACTAGATCGTGTGGGCATC-----CAGGCTGCTCTCTGAGGATACCTC 518
Db 436 ValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPhe 455
QY 519 TCCCTGATGCATTAAATGCAGAGGTGATATCTTCAGGGTGTCTTATGCTGGGGCCCCCA 578
Db 456 LysTyrLys-----MetValAlaIleAlaGlyAlaPro 466
QY 579 GTCACTCTGTGGATCTTTCTATGATACAGGATACAGGAACTGTTATATGGGTCACTTCGAC 638
Db 467 ValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAsp 486
QY 639 CAGAAATGAACAGGCTATTACTTAGGATCTGTGGCATCAGCAGGAAAGTTCCCTCT 698
Db 487 GlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSer 506
```





```
; NUMBER OF SEQ-ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-415-122-4

Alignment Scores:
Pred. No.: 1,13e-106 Length: 869
Score: 1094.50 Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1% Indels: 49
DB: 4 Gaps: 1

US-10-825-632-4 (1-1197) x US-10-415-122-4 (1-869)

QY 3 TTTGAAGGACCAAGACTCCCTTTAGAGCATCAGCTGTAGTAGTCAGTAAAT 62
Db 521 PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluSer 540
QY 63 CCTGGAGAGGTGACAGGCTGACTGACCGTGCTACTCACATTCCTGCTGCATCAGTCAG 122
Db 541 AlaGlyGluLeuValArgLeuThrThrLeuGlyPheSerHisSerCysSerMetSerGln 560
QY 123 CACTGTGACTCTTTTAAAGTAAGTATAGTAAACCAAGAAGTCCACACTGTGTCCCTT 182
Db 561 SerPheAspMetPheValSerHisTyrSerValSerThrProCysValHisVal 580
QY 183 TACAGCTATCAAGTCTCTGAAGATGACCCCACTTCGAAACAAAGGAATTTTGGGCCACC 242
Db 581 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAlaSer 600
QY 243 ATTTGGATTACGACGCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGA 302
Db 601 MetMetGluAlaAlaSerCysProProAspTyrValProProGluLeuPheHisPheHis 620
QY 303 AGTACTACTGGATTTACATTGTATGGGATGCTCTCAAGCCCTCATGATCTACAGCCCTGGA 362
Db 621 ThrArgAlaAspValGlnLeuTyrGlyMetIleTyrLysProHisThrLeuGlnProGly 640
QY 363 AAGAAATATCTACTGTGCTGTTCATATATGTGTCTCTCAG----- 404
Db 641 ArgLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 660
QY 404 ----- 404
Db 661 SerPheLysGlyIleLysTyrLeuArgLeuAenThrLeuAlaSerLeuGlyTyrAlaVal 680
QY 404 ----- 404
Db 681 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuHisPheGluGlyAlaLeuLys 700
QY 405 -----GGTCAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGTTCT 455
Db 701 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnTyrValAlaGlu 720
QY 456 CGATATGATTCTTACATCTAGATCTGTGGGCATCCACGGCTGCTCTATGGAGGATAC 515
Db 721 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyIleTrpSerTyrGlyGlyPhe 740
QY 516 CTCTCCCTGATGCGATTAAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
Db 741 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 760
QY 576 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACCT 635
Db 761 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 780
QY 636 GACCAGATGAACAGGGCTATTACTAGTCTGTGGCCATGCAAGCAGAGAAAGTCCCC 695
Db 781 GluAsnAsnGlnGlnGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 800

696 TCTGAACCAAAATCCTTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTCGA 755
801 AsnGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPhePhe 820
756 CATACCACTATATTACTGAGTCTTTTAGTGGGCTGGAAGCCCATATGATTACAGATC 815
821 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 840
816 TATCTCTCAGGAGACACACATAGAGTTCTCTGAATCGGAGAACATTATGAACCTGCAT 875
841 TyrProAsnGluArgHisSerIleArgCysArgGluSerGlyGluHisTyrGluValThr 860
876 CTTTGTGACTACCTTCAAGAAAACCTT 902
861 LeuLeuHisPheLeuGlnGluHisLeu 869

Search completed: May 2, 2006, 03:08:34
Job time : 190.411 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Biocolloration Ltd.

OM nucleic - protein search using **beam\_plus\_n2p** model

Run on: May 2, 2006, 00:41:48 ; Search time 36.3435 Seconds  
(without alignments)  
2894.257 Million cell updates/sec

Title: US-10-825-632-4  
Perfect score: 2140  
Sequence: 1 attttgaaggcaccacaaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-O=/abs/ABSSWEB.spool/US10825632/runat\_01052006.105944.3214/app query.fasta.1  
-DB=A Geneseg -QFWT=fascan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOFCI=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p  
-USER=US10825632 @CGN 1.1.605 @runat\_01052006.105944.3214 -NCPU=6 -ICPU=3  
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseg 21.\*  
1: Genesegp1980s.\*  
2: Genesegp1990s.\*  
3: Genesegp2000s.\*  
4: Genesegp2001s.\*  
5: Genesegp2002s.\*  
6: Genesegp2003as.\*  
7: Genesegp2003bs.\*  
8: Genesegp2004s.\*  
9: Genesegp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1680	78.5	310	4 AAB47188	Aab47188 Human DPP
2	1680	78.5	310	5 ABB08994	Abb08994 Human dip
3	1680	78.5	310	7 ADR27858	Adr27858 Human dip
4	1645.5	76.9	587	5 ADR41398	Adr41398 Human CD-
5	1645.5	76.9	824	6 AABU92030	Aabu92030 Human pro
6	1645.5	76.9	882	4 AAB47187	Aab47187 Human DPP
7	1645.5	76.9	882	5 ABB61591	Abb61591 Human DPP
8	1645.5	76.9	882	5 AAG78415	Aag78415 Amino aci
9	1645.5	76.9	882	5 AAE24170	Aae24170 Human dip

10	1645.5	76.9	882	5 AAU74749	Aau74749 Human pro
11	1645.5	76.9	882	5 ADI17086	Adi17086 Human NOV
12	1645.5	76.9	882	6 ABU07720	Abu07720 Human ser
13	1605.5	75.0	883	5 ADI17085	Adi17085 Murine NO
14	1407	65.7	493	7 ADE78977	Ade78977 Human pro
15	1407	65.7	831	6 ABU92026	Abu92026 Human pro
16	1401	65.5	632	4 AAB93585	Aab93585 Human pro
17	1377.5	64.4	465	4 AAB47189	Aab47189 Human DPP
18	1377.5	64.4	724	5 ABB97362	Abb97362 Novel hum
19	1377.5	64.4	782	5 ABB97361	Abb97361 Novel hum
20	1372	64.1	516	6 ABU92029	Abu92029 Human pro
21	1372	64.1	580	5 AAE14337	Aae14337 Human pro
22	1337	62.5	738	8 ADT04044	Adt04044 Human pro
23	1094.5	51.1	869	5 AAE24169	Aae24169 Alternati
24	1093.5	51.1	497	5 AAU96169	Aau96169 Human sec
25	1093.5	51.1	497	5 ABG64845	Abg64845 Human alb
26	1093.5	51.1	497	8 ADL78112	Adl78112 Albumin f
27	1093.5	51.1	597	7 ADE08485	Ade08485 Novel pro
28	1093.5	51.1	689	5 ADI17084	Adi17084 Human NOV
29	1093.5	51.1	830	5 AAE24171	Aae24171 Human dip
30	1093.5	51.1	863	5 ABG61592	Abg61592 Human DPP
31	1093.5	51.1	863	5 ADI17083	Adi17083 Human NOV
32	1093.5	51.1	892	5 ABG61602	Abg61602 Human DPR
33	1093.5	51.1	892	5 ABG61604	Abg61604 Human DPR
34	1093.5	51.1	892	5 ABB98134	Abb98134 Human PWM
35	1093.5	51.1	892	8 ADS10951	Ads10951 Human the
36	1093.5	51.1	969	5 AAE24168	Aae24168 Human dip
37	1089.5	50.9	497	5 AAU96192	Aau96192 Human sec
38	1089.5	50.9	497	5 ABG64844	Abg64844 Human alb
39	1089.5	50.9	497	8 ADL78111	Adl78111 Albumin f
40	1089.5	50.9	581	5 ADR41521	Adr41521 Human CD-
41	1086.5	50.8	360	4 AAB47190	Aab47190 Human DPP
42	1086.5	50.8	755	8 AADQ67811	Aadq67811 Novel hum
43	1086.5	50.8	879	5 ABG61607	Abg61607 Human DPR
44	1086.5	50.8	879	5 ABG61608	Abg61608 Human DPR
45	1086.5	50.8	879	8 ADS10953	Ads10953 Human the

ALIGNMENTS

RESULT 1  
AAB47188  
ID AAB47188 standard; protein; 310 AA.  
XX  
XX  
AC AAB47188;  
XX  
DT 29-JUN-2001 (first entry)  
XX  
DE Human DPP8 524Phe-657Gln+707Gly-882Ile.  
XX

Human; dipeptidyl aminopeptidase; DPP8; prollyl oligopeptidase;  
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;  
KW growth hormone deficiency; glucose level; mucosal regeneration;  
KW non-insulin dependent diabetes mellitus; glucose intolerance;  
KW immunosuppression.

XX Homo sapiens.  
XX  
XX WO200119866-A1.  
XX  
XX 22-MAR-2001.  
XX

XX 11-SEP-2000; 2000WO-AU001085.  
XX  
XX 10-SEP-1999; 99AU-00002762.  
XX 18-FEB-2000; 2000AU-00005709.  
XX (UNSY ) UNIV SYDNEY.  
XX  
XX Abbott CA, Gorell MD;  
XX  
XX WPI; 2001-281520/29.  
XX  
XX N-PSDB; AAC85695.



Pred. No.:	1.88e-184	Length:	310
Score:	1680.00	Matches:	310
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	78.5%	Indels:	0
DB:	5	Gaps:	0
US-10-825-632-4 (1-1197) x ABB08994 (1-310)			
QY	3	TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGTACGTAGTCAGTTACGTAAT	62
Db	1	PheGluGlyThrIysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	20
QY	63	CCTGGACAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTCTGCTGCATCAGTCAG	122
Db	21	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln	40
QY	123	CACGTGCTCTTTTATAGTATAGTAACTACACAGAAATCCACACTGTGTCTCCCTT	182
Db	41	HisCysAspPhePheIleSerIysTyrSerAsnGlnLysAsnProHisCysValSerLeu	60
QY	183	TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAACAAAGGAATTTTGGCCACC	242
Db	61	TyrIysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr	80
QY	243	ATTTGGATTACAGAGTCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGA	302
Db	81	IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu	100
QY	303	AGTACTACTGGATTACATTGATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA	362
Db	101	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	120
QY	363	AAGAAATATCTACTGTCTCATATATGTGTCTCTCAGGTCAAATAGAAATGAC	422
Db	121	LysLysTyrProThrValLeuPheIleTyrGlyProGlnGlyGlnIleGluIleAsp	140
QY	423	GATCAGTGAAGGACTCCAAATATCTAGCTCTCCATATGATTTTCATTGACTAGTCGT	482
Db	141	AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg	160
QY	483	GTGGCATCCACCGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAAATCGACAGG	542
Db	161	ValGlyIleHisGlyTrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArg	180
QY	543	TCAGATATCTTCAGGTTCTATTGCTGGGGCCCCAGTCCTCTGTGATCTTCTATGAT	602
Db	181	SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp	200
QY	603	ACAGGATACCGGAACGTTATATGGGTACCTCTGACAGAAATGAACAGGGCTATTACTTA	662
Db	201	ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu	220
QY	663	GGATCTGTGCGCATGCAACGAGAAAGTTCCCTCTGAACCAATCGTTTACTCTCTTA	722
Db	221	GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu	240
QY	723	CATGTTTCTCTGATCAGAAATGTCACATTTTGACATACACAGTATATTACTGAGTTT	782
Db	241	HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu	260
QY	783	GTGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACAGCATAAGA	842
Db	261	ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg	280
QY	843	GTTCCTCAATCGGGAGAACATTATGAACTGCATCTTTTGCATCTACCTTCAAGAAACCTT	902
Db	281	ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGlnAsnLeu	300
QY	903	GGATCAGTATTCTCTCTTAAAGTGATA	932
Db	301	GlySerArgIleAlaLeuLysValIle	310
RESULT 3			
ADD27858			
ID	ADD27858	standard; protein; 310 AA.	
XX	AC	ADD27858;	
XX	DT	15-JAN-2004 (first entry)	
XX	DE	Human dipeptidyl peptidase 8.	
XX	KW	Mucosal inflammation; rhinitis; sinusitis; exopeptidase; substance P; SP; neurokinin 1 receptor; NK1 receptor; allergy; asthma; antiasthmatic; human; dipeptidyl peptidase 8; enzyme.	
XX	OS	Homo sapiens.	
XX	PN	US2003165489-A1.	
XX	PD	04-SEP-2003.	
XX	PF	27-NOV-2001; 2001US-00993959.	
XX	PR	28-FEB-2001; 2001US-00794236.	
XX	PA	(BMRA-) BMRA CORP BV.	
XX	PI	Grouzmann E, Lacroix J, Monod M;	
XX	DR	WPI; 2003-811386/76.	
XX	PT	Treatment of patient for mucosal inflammation associated with rhinitis and/or sinusitis involves intranasally administering peptidase that cleaves at Xaa-Pro sequences or agent inhibiting binding of Sp to neurokinin 1 receptor.	
XX	PS	Disclosure; SEQ ID NO 4; 14pp; English.	
XX	CC	The present invention relates to a method of treating a patient for mucosal inflammation associated with rhinitis and/or sinusitis. The method comprises intranasally administering to the patient a peptidase that cleaves at Xaa-Pro sequences or an agent that inhibits the binding of substance P (SP) to the neurokinin-1 (NK1) receptor. The peptidase is an exopeptidase, preferably selected from human dipeptidyl peptidase IV (DPPIV), human quiescent cell proline dipeptidase, human dipeptidyl peptidase 8, or human attractin. The method is useful for treating a patient for mucosal inflammation associated with rhinitis and/or sinusitis which are the result of allergies or asthma. The invention provides an effective treatment of the inflammation associated with both rhinitis and sinusitis. The present sequence represents human dipeptidyl peptidase 8.	
XX	SQ	Sequence 310 AA;	
Alignment Scores:			
Pred. No.:	1.88e-184	Length:	310
Score:	1680.00	Matches:	310
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	78.5%	Indels:	0
DB:	7	Gaps:	0
US-10-825-632-4 (1-1197) x ADD27858 (1-310)			
QY	3	TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGTACGTAGTCAGTTACGTAAT	62
Db	1	PheGluGlyThrIysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	20
QY	63	CCTGGACAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTCTGCTGCATCAGTCAG	122
Db	21	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln	40
QY	123	CACGTGCTCTTTTATAGTATAGTAACTAGTAACTACACAGAAATCCACACTGTGTCTCCCTT	182

Db 41 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 60  
QY 183 TACAAGCTATCAAGTCCTGAAGATGACCCAACTTGGCAAAACAAGAAATTTGGGCGACC 242  
Db 61 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheThrAlaThr 80  
QY 243 ATTTTGGAATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302  
Db 81 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 100  
QY 303 AGTACTACTGGATTTTACATTTGATGGATGCTCTACAAAGCTCATGATCTACAGCCTGGA 362  
Db 101 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 120  
QY 363 AAGAAATATCTACTGCTGCTTCATATATGCTGCTCAGGTCAGGTCAGAAATAGAC 422  
Db 121 LysLysTyrProThrValLeuPheIleTyrGlyProGlnGlyGlnIleGluLeuAsp 140  
QY 423 GATCAGGTCGAAGGACTCGATATCTAGCTTCTCGATATGATTTTCATTGATTCGT 482  
Db 141 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 160  
QY 483 GTGGGATCCAGCGCTGCTTATGAGGATGATCTCTCCCTGATGGCATTAATGACAGG 542  
Db 161 ValGlyIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 180  
QY 543 TCAGATATCTCAGGCTGCTATCTCTGGGCCCCAGTCACTCTGTGGATCTTCTATGAT 602  
Db 181 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 200  
QY 603 ACAGGATACACGGAAGCTTATATGGTCCACCTGACCAAGATGACAGGCTTATCTTA 662  
Db 201 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu 220  
QY 663 GGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTA 722  
Db 221 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeu 240  
QY 723 CATGTTTCTCTGGATGAGAACTTCCATTTTGCATACCACTGATATATTACTGATTTT 782  
Db 241 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 260  
QY 783 GTGAGGCTGGAAAGCCATGATTTACAGATCTATCTCAGAGGAGACACAGCATAAGA 842  
Db 261 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 280  
QY 843 GTTCTCGATCGGAGACATTTATCACTGATCTTTTGCATCTTTTGCATCTTCAAGAAACCTT 902  
Db 281 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 300  
QY 903 GATCAGCTATTGCTGCTCTAAAGTGATA 932  
Db 301 GlySerArgIleAlaAlaLeuLysValIle 310  
RESULT 4  
ID ADR41398  
AC ADR41398 standard; protein; 587 AA.  
XX ADR41398;  
XX  
DT 07-OCT-2004 (first entry)  
DE Human CD-like molecule HDQES11, SEQ ID NO:197.  
XX  
XX Human; CD-like molecule; cluster of differentiation; diagnosis;  
KW prevention; immune disorder; immunodeficiency; autoimmune disorder;  
KW blood-related disorder; haematological disorder; haemostatic disorder;  
KW thrombolytic disorder; hyperproliferative disorder; cancer; tumour;  
KW apoptotic disorder; cardiovascular disorder; respiratory disorder;  
KW angiogenic disorder; neovascularisation; neurological disorder;  
KW endocrine disorder; reproductive system disorder; infectious disease;  
KW gastrointestinal disorder; drug screening; tissue regeneration;  
KW chemotaxis; gene therapy; antibody therapy; drug targeting;

KW chromosome mapping; forensic analysis; immunophenotyping; cytostatic;  
KW haemostatic; tranquiliser; vulnerary; anti-inflammatory; nephrotropic;  
KW cardiant; anti-allergic; anti-HIV; antirheumatic; antiarthritic;  
KW antipsoriatic; immunosuppressive; vasotropic; neuroprotective;  
KW antithyroid; thyrominetic; gynaecological; virucide; hepatotropic;  
KW antibacterial; dermatological; chromosome 15q22.  
OS Homo sapiens.  
XX  
XX WO200226930-A2.  
XX  
PD 04-APR-2002.  
XX  
PF 25-SEP-2001; 2001WO-US029838.  
XX  
PR 26-SEP-2000; 2000US-0235484P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Birse CE;  
FI WPI, 2002-405050/43.  
DR N-PSDB; ADR41222.  
XX  
PT Novel polynucleotides and polypeptides useful for treating, preventing or  
PT ameliorating cardiovascular, renal, neurovascular, and autoimmune  
PT disorders.  
XX  
PS Claim 11; SEQ ID NO 197; 1243pp; English.  
XX  
CC The invention relates to 167 novel human CD (cluster of differentiation)-  
CC like molecules (ADR41388-ADR41563) and to cDNAs encoding them (seqid:1)-  
XX  
SQ Sequence 587 AA;  
Alignment Scores:  
Pred. No.: 2,53e-180 Length: 587  
Score: 1645.50 Matches: 310  
Percent Similarity: 86.4% Conservative: 0  
Best Local Similarity: 86.4% Mismatches: 0  
Query Match: 76.9% Indels: 49  
DB: 5 Gaps: 1  
US-10-825-632-4 (1-1197) x ADR41398 (1-587)  
QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTCAGTTACGTAAAT 62  
Db 229 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 248  
QY 63 CCTGGAGAGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGCTCATCAGTCAG 122  
Db 249 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 268  
QY 123 CACTGTGACTTCTTTATAGTATAGTATAGTACAGAGAGAAATCCACACTGTGTCTCCTT 182  
Db 269 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 288  
QY 183 TACAGCTATCAAGTCTCTGAAAGATGCCCACTTCGCAAAACAAGAAATTTGGGCCACC 242  
Db 289 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheThrAlaThr 308  
QY 243 ATTTTGGAATTCAGCAGGTCCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302  
Db 309 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 328  
QY 303 AGTACTACTGGATTTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 362  
Db 329 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 348  
QY 363 AAGAAATATCTACTGCTGCTGTTTCATATATGCTGCTCTCTAG----- 404  
Db 349 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn 368



```
QY 404 ----- 404
Db 369 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 388
QY 404 ----- 404
Db 389 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 408
QY 405 -----GTCGAATAGAAATTGCGATCGAGTGGAGGAGCTCCATATCTAGCTTCT 455
Db 409 TyrLysMetGluGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 428
QY 456 CGATATGATTTCACTTAGATCGTGCGGCATCCACGGCTGGTCTCTATGGAGGATAC 515
Db 429 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheTyrGlyGlyTyr 448
QY 516 CTCTCCCTGATGCGATTAATGACAGAGTCAGATATCTTCAGGGTTCGTATTGCTGGGGCC 575
Db 449 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 468
QY 576 CCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCT 635
Db 469 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 488
QY 636 GACCAAGATGACAGGGCTATTACTTAGATCTGTGGCCATCCAAAGCAGAAAGTTCCCC 695
Db 489 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 508
QY 696 TCTGAACCAATCGTTTACTGCTTACATGCTTACATGCTTCCTGGATGAGAAATGTCATTGGCA 755
Db 509 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 528
QY 756 CATACCAAGTATATTACTAGTTTGTAGTGAGGGCTGGAAGCCATATGATTTACAGATC 815
Db 529 HisthrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 548
QY 816 TATCTCAGGAGACACACATAAGAGTTCTCGAATCGGAGAGACATTATGCACTGCAT 875
Db 549 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 568
QY 876 CTTTGTGCACTACCTTCAAGAAACCTTTGGATCAGCTATTGCTGCTCTAAAGTGATA 932
Db 569 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 587

RESULT 5
ABU92030
ID ABU92030 standard; protein; 824 AA.
AC ABU92030;
XX
XX
XX 15-JUL-2003 (first entry)
XX Human protein modification and maintenance molecule-10 (PMMW-10).
XX Human; protein modification and maintenance molecule; PMMW; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; anti-allergic; vulnerary;
KW antiinflammatory; thyromimetic.
XX Homo sapiens.
OS
XX WO2003031939-A2.
PN
XX 17-APR-2003.
PD
XX 11-OCT-2002; 2002MO-US032850.
PF
XX 12-OCT-2001; 2001US-0329689P.
PR
XX 25-OCT-2001; 2001US-0335703P.
PR
```

```
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY,
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hafalia AUA, Chawla NK, Lehr-Mason PW, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX WPI: 2003-430274/40.
DR N-PSDB; ACA92425.
XX
XX New human protein modification and maintenance molecules (PMMW), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMW expression e.g. cancer, atherosclerosis, or
PT infections.
XX Claim 1; Page 249-251; 31pp; English.
XX
XX The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMW), and the polynucleotide
CC sequences encoding them. A total of 40 PMMW polypeptides (designated PMMW
CC -1 to PMMW-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMW in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMW are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), gastrointestinal disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMW polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMMW polypeptides of the invention
XX
XX Sequence 824 AA;

Alignment Scores:
Pred. No.: 3e-180 Length: 824
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 6 Gaps: 1

US-10-825-632-4 (1-1197) x ABU92030 (1-824)
QY 3 TTTGAGGACCAAGAAGCTCCCTTTAGACATCACTCTGACGTAGTACGTACGTAAT 62
Db 466 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 485
QY 63 CTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCATCTTCTGTCATCATGTCAG 122
Db 486 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 505
QY 123 CACTGTGACTTCTTTTATAAGTAAAGTATAGTAACCAAGAATCCACACTGTGTGCCCTT 182
Db 506 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 525
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCACAAACAAAGAAATTTTGGCCACC 242
Db 526 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 545
QY 243 ATTTTGGATTCAGCAGGTCTCTTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 302
Db 546 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 565
```

QY	303	AGTACTACCTGGATTACATTGATGGGATGCTCTACAGGCTCATGATCTACAGCTCGA	362	FT	Active-site	/note= "Forms part of Ser-Asp-His catalytic triad"
Db	566	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	585	FT	Active-site	/note= "Forms part of Ser-Asp-His catalytic triad"
QY	363	AAGAAATATCTACTGTCTGCTTCATATATGCTGCTCTCAG	404	FT	Active-site	/note= "Forms part of Ser-Asp-His catalytic triad"
Db	586	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn	605	PN	WO200119866-A1.	
QY	404	-----	404	PD	22-MAR-2001.	
Db	606	ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal	625	PF	11-SEP-2000; 2000WO-AU001085.	
QY	404	-----	404	PR	10-SEP-1999; 99AU-00002762.	
Db	626	ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	645	PR	18-FEB-2000; 2000AU-00005709.	
QY	405	-----GGTCAATGAAATTCACGATCAGGTGGAAGACTCCCAATATCTAGCTTCT	455	PA	(UNSY ) UNIV SYDNEY.	
Db	646	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	665	PI	Abbott CA, Gorell MD;	
QY	456	CGATATGATTTCATTCAGCTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATAC	515	DR	WPI; 2001-281520/29.	
Db	666	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr	685	XX	N-PSDB; AAC85694.	
QY	516	CTCTCCCTGATGATATGACGAGCTCAGATATCTTCAGGTTGCTATTGCTGGGCC	575	XX	New human dipeptidyl aminopeptidase (DPP8) useful for cleaving	
Db	686	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	705	XX	substrates, identifying inhibitors of DPP8 catalytic activity which have	
QY	576	CCAGTCACCTGTGGATCTTCTATGATACAGGATACACGAGCTTATGCTGCTACCCCT	635	XX	therapeutic uses, and for detecting activated T cells.	
Db	706	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro	725	XX	Claim 1; Fig 2; 78pp; English.	
QY	636	GACAGAATGAACAGGGCTATTACTAGGATCTGTGCCATGTGCAAGCAGAAAGTCCCC	695	XX	This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has	
Db	726	AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro	745	XX	substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.	
QY	696	TCGAACCAAACTCGTTACTGCTTACATGCTTCTGATGATGATGCTTCTGATGATGCT	755	XX	Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase,	
Db	746	SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla	765	XX	because it is capable of hydrolysing the peptide bond C-terminal to Pro	
QY	756	CATACAGTATATCTAGTGTGTTTGTAGTGAGGGCTGGAAGCCATATGATTTACAGATC	815	XX	in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is	
Db	766	HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle	785	XX	useful for cleaving a substrate, and for detecting an activated T cell	
QY	816	TATCCTCAGGAGACACAGCATAAGAGTTCTGATCGGAGAACATTATGAAGTCAT	875	XX	which involves measuring the level of DPP8 gene expression in a T cell.	
Db	786	TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis	805	XX	The level of DPP8 expression is detected by detecting the amount of DPP8	
QY	876	CTTTGCACTACCTTCAGAAACCTTGATCAGTATGCTGCTCTAAAAGTGATA	932	XX	RNA in the cell. It is also useful for identifying a molecule capable of	
Db	806	LeuLeuHisTyrLeuGlnLeuGlnLeuGlySerArgIleAlaLeuLysValIle	824	XX	inhibiting the cleavage of the substrate by DPP8. Molecules identified as	
RESULT 6				XX	CC growth hormone deficiency, lowering glucose levels in non-insulin	
AAB47187				XX	dependent diabetes mellitus and other disorders involving glucose	
ID	AAB47187	standard; protein; 882 AA.		XX	intolerance, enhancing mucosal regeneration and as immunosuppressants	
XX	AC	AAB47187;		XX	Sequence 882 AA;	
XX	DT	29-JUN-2001 (first entry)		XX	Alignment Scores:	
XX	DE	Human DPP8.		XX	Pred. No.:	3.1e-180
XX	XX			XX	Score:	Length: 882
KW	KW	Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;		XX	Percent Similarity:	Matches: 310
KW	KW	dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;		XX	Best Local Similarity:	Conservative: 0
KW	KW	growth hormone deficiency; glucose level; mucosal regeneration;		XX	Query Match:	Mismatches: 0
KW	KW	non-insulin dependent diabetes mellitus; glucose intolerance;		XX	DB:	Indels: 49
XX	OS	immunosuppression.		XX		Gaps: 1
XX	OS	Homo sapiens.		XX	US-10-825-632-4 (1-1197) x AAB47187 (1-882)	
XX	Key	Location/Qualifiers		QY	3	TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAAAT
XX	Key	739		Db	524	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn
FT	Active-site			QY	63	CCTGGAGAGGTGCAAGGCTGACTGACCGTGCTACTCAGCTTCTGTCAGTCAGTCAG
				Db	544	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln
				QY	123	CAGTGTGACTCTTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTAT
				Db	564	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu
				QY	183	TACAAGCTATCAAGTCTCTGAAGTATGACCCCACTTGCAGAAACAGAAATTTGGCCAC
				Db	594	TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr
				QY	243	ATTTTGAATTCAGAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA

Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623  
Qy 303 AGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362  
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
Qy 363 AGAATATCTCTACTGCTGCTTCATATATATGCTGCTCTCAG-----404  
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663  
Qy 404 -----404  
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
Qy 404 -----404  
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
Qy 405 -----GGTCAATAGAAATTGACCATCAGGTGGAGGACTCCAAATATCTAGCTTCT 455  
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
Qy 456 CGATATGATTCTATGACTTACATCGTGTGGGCATCCACGGCTGCTCTATGGAGATAC 515  
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743  
Qy 516 CTCTCCTCATGGCATTATGACAGAGGTACAGATATCTTCAGGGTGTCTATGTGGGGCC 575  
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
Qy 576 CAGTCACCTCTGTGATCTTCTATGATACAGGATACAGGAACTTATATGGTCCACCT 635  
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783  
Qy 636 GACCAGAAAGACAGGGCTATTACTTGGATCTGTGGCCATGCCAAGCAGAAAGTCCCC 695  
Db 784 AspGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803  
Qy 696 TCTGAACCAAACTGTTACTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTTGA 755  
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823  
Qy 756 CATACCATATATCTAGTTTTTATGAGGGCTGGAAGCCATATGATTTACAGATC 815  
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843  
Qy 816 TATCTCAGAGACACAGCATAGACTTCTGAATCGGGAGAACATTATGAATCCAT 875  
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863  
Qy 876 CTTTTGCATCTACCTCAAGAAACCTTGGATCAGCTATTGCTCTTAAAGTGATA 932  
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882

RESULT 7  
ABG61591  
ID ABG61591 standard; protein; 882 AA.  
AC ABG61591;  
XX  
DT 12-AUG-2003 (first entry)  
DE Human DPPIV related serine protease DPP-1.  
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPP;  
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;  
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;  
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;  
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;  
KW dyskinesia; reproductive disorder; inflammatory disorder;  
KW metabolic disorder.  
XX Homo sapiens.

XX WO200231134-A2.  
PN 18-APR-2002.  
XX 12-OCT-2001; 2001WO-US031874.  
PF 12-OCT-2000; 2000US-0240117P.  
PR (FERR ) FERRING BV.  
PA  
XX Qi S, Akineanya KO, Riviere PJ, Junien J;  
PI WPI; 2002-444178/47.  
XX N-PSDB; ABK83322.  
DR  
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding  
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and  
PT viral infections, cancers, allergies, neurological disorders, or pain.  
XX Claim 17; Fig 1; 113pp; English.  
XX The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins  
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic  
CC acids encoding them are useful for treating infections such as fungal,  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591  
CC -ABG61612 represent human DPRP proteins  
XX  
SQ Sequence 882 AA;  
Alignment Scores:  
Pred. No.: 3.1e-180 Length: 882  
Score: 1645.50 Matches: 310  
Percent Similarity: 86.4% Conservatives: 0  
Best Local Similarity: 86.4% Mismatches: 0  
Query Match: 76.9% Indels: 49  
DB: 5 Gaps: 1  
US-10-825-632-4 (1-1197) x ABG61591 (1-882)  
Qy 3 TTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGTGTAGTGTAGTGTAGTGTAGT 62  
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543  
Qy 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTGTGTGTGTGTGT 122  
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563  
Qy 123 CACTGTGACTTCTTTATAGTAAAGTATAGTAAACCAAGAAATCCACACTGTGTGTGTGTGT 182  
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
Qy 183 TACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGAAATTTGGGCCACC 242  
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
Qy 243 ATTTTGGATTCAGAGGTCTCTCTCTGATATCTCTCCAGAAATTTCTCTTTTGA 302  
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623  
Qy 303 AGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362  
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643

```
QY 363 AAGAAATATCTACTGCTGCTGTTTCATATATGCTGCTCTCAG----- 404
Db 644 LybelytyrProthrValleuPheiletyrGlyglyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
Db 664 ArgPheLyglyVallytyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 ----- 404
Db 684 valValIleAspAsnArgGlySerCysHisArgGlyLeuLySpheGluGlyAlaPheLyS 703
QY 405 -----GTCAAATAGAAATGACCATCAGGTGGAGACTCCCAATATCTAGTTCT 455
Db 704 TyrLyMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CGATATGATTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTATGGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
QY 516 CTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTGTGCTATTGCTGGGCG 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CAGTCACNCTGTGATCTTCTATGATACAGGATACAGGAAGCTTATATGGTCCACCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAAGTACAGGGCTATTACTAGATCTGTGGCCATCGACGAGAAAGTTCCCG 695
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLySphePro 803
QY 696 TCTGAACCAATCGTTTACTGCTCTTACATCGTTTCTGAGTGAAGATGCCATTTGCA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACAGPATATTACTGAGTTTTTTAGTGAGGCTCGAAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLySpProTyrAspLeuGlnIle 843
QY 816 TATCCTCAGGAGACACAGCATAGAGTTCTTGAATCGGGAGAACATTTATGAATGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTTGCACCTCTTCAAGAAACCTTGGATCAGTATTGCTCTCTAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLySValIle 882

RESULT 8
ID AAG78415
XX AAG78415 standard; protein; 882 AA.
XX AAG78415;
XX 12-APR-2002 (first entry)
DE Amino acid sequence of 21953 human prollyl oligopeptidase.
XX 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease.
XX Homo sapiens.
XX WO2001/9473-A2.
XX 25-OCT-2001.
XX 11-APR-2001; 2001WO-US040483.
XX
```

```
PR 18-APR-2000; 2000US-0197508P.
XX (MILL-) MILLENNIUM PHARM INC.
XX PI Meyers RA, Williamson M;
XX WPI; 2002-034353/04.
DR N-PSDB; AAH99934.
XX New polypeptides 21953, member of human prollyl oligopeptidase family,
PT useful as diagnostic targets and therapeutic agents for controlling
PT cancer, lymphoma and leukemia.
PS Claim 1; Page 102-103; 121pp; English.
CC This invention relates to an isolated 21953 human prollyl oligopeptidase.
CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
CC antithyroid, dermatological, antiepileptic, antiasthmatic,
CC ophthalmological, antiinflammatory, nootropic, antiparkinsonian,
CC anticonvulsant, gynaecological, vasotropic, antianginal, cardiac,
CC antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC gene therapy, expression or activity of 21953 protein modulator, it is
CC useful for identifying a compound which binds to it and can be used in
CC preventing, treating or detecting a cellular proliferative or
CC differentiative disorder. The 21953 molecules can act as novel diagnostic
CC targets and therapeutic agents for controlling disorders associated with
CC the aberrant activity or degradation of peptide hormones e.g., cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, demyelinating diseases, vascular disorders and
CC metabolism or pain disorders. This sequence represents the amino acid
CC sequence of 21953 human prollyl oligopeptidase
XX
SQ Sequence 882 AA;

Alignment Scores:
Pred. No.: 3,1e-180 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 5 Gaps: 1

US-10-825-632-4 (1-1197) x AAG78415 (1-882)
QY 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGTTAGTAAAT 62
Db 524 PheGluGlyThrLyAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 63 CTTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGCTAGTCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysSileSerGln 563
QY 123 CATGTGACTCTCTTATATAAGTATATAGTAAACGAGAAGATCCACACTGTGTGTCCTT 182
Db 564 HisCysAspPhePheIleSerLySerAenGlnLyAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTTGAAGATGACCCCACTTGTCAAAACAAAGGAATTTTGGGCCACC 242
Db 584 TyrLyLeuSerSerProGluAspAspProThrCysLyThrLySLeuGluPheTrpAlaThr 603
QY 243 ATTTTGGATTACAGAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGGATTTTACATTGTATGGGATGCTCTACAGGCTCTATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLySProHisAspLeuGlnProGly 643
```

```
QY 363 AAGAAATATCTACTGCTGCTGTTTCATATATATGCTGCTCTCAG-----404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsn 663
QY 404 -----404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 693
QY 404 -----404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----404
Db 704 TyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 456 CGATATGATTTCATGACTTACATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743
QY 516 CTCTCCCTGATGGCATTAAATGACAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACGTTATATGGGTCACTT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
QY 636 GACCAGAATGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTCCCC 695
Db 784 AbpGlnAsnGlnGlnGlyTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGAACCAAACTGTTACTGCTTACATGCTTTCCTGGATGAGAAATGTCATTTGCA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACAGTATATTACTGAGTTTTTATGAGGGCTGGAAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
QY 816 TATCCTCAGGAGACACAGCATAGAGTCTCTGAATCGGGAGAACATTTAGCTGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
QY 876 CTTTTCACATCTACTCTCAAGAAACCTTGGATCAGCTATTGCTCTTAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882

RESULT 9
AAE24170
ID AAE24170 standard; protein; 882 AA.
XX
AC AAE24170;
XX
XX
DT 23-SEP-2002 (first entry)
XX
DE Human dipeptidyl peptidase 8 (DPP8) protein.
XX
KW Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW antiviral; enzyme.
XX
OS Homo sapiens.
XX
PN WO200234900-A1.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-AU001388.
XX
PR 27-OCT-2000; 2000AU-00001078.
XX
```

```
PA (UNSY ) UNIV SYDNEY.
XX
XX PI Abbott CA, Gorrell MD;
XX
XX WPI; 2002-454646/48.
DR N-P8DB; RAD38956.
DR
XX
XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
XX of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
XX type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
XX infection.
XX
XX Example; Fig 1; 91pp; English.
PS
XX
XX The present invention relates to dipeptidyl peptidase (DPP) proteins and
XX polynucleotides encoding such proteins. The DPP peptides are useful for
XX screening inhibitors of DPP catalytic activity. The inhibitors are useful
XX for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
XX rejection and HIV (human immuno deficiency virus) infection. The present
XX sequence is human DPP8 protein
XX
XX Sequence 882 AA;
SQ
```

```
Alignment Scores:
Pred. No.: 3.1e-180 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 5 Gaps: 1
```

US-10-825-632-4 (1-1197) x AAE24170 (1-882)

```
QY 3 TTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTACGTAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 63 CCTGGAGAGTGACAGGCTGACTGACCGTGGCTACTCACATTCCTTGTGTCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 123 CACTGTGACTCTTTTATATAGTATAGTATACACAGAGAAATCCACACTGTGTCTCCCTT 182
Db 564 HisCysAspPhePheIleSerTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCACTTCGAAACAAAGAAATTTGGSCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTGGATTTCAGCAGGCTCCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCCCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 363 AAGAAATATCTACTGCTGCTTTTCATATATGTTGTTGCTCTCAG-----404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn 663
QY 404 -----404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 404 -----404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----404
Db 704 TyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
```

```
QY 456 CGATATCATTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATAC 515
Db 724 ArgyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrglyGlyTy 743
QY 516 CTCTCCCTGATGCGATTAAAGCAGAGGTCAGATATCTTCAGGTTGCTATTGTGGGGCC 575
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 576 CCAGTCACCTCTGTGGATCTCTATCATACAGGATACACGGAACTTATATGGGTCCACCT 635
Db 764 ProValThrLeuTrpIlePheTyArgThrGlyTyThrGluArgTyMetGlyHisPro 783
QY 636 GACCAGAAATGAACAGGGCTATTACTAGGATCTGTGGCATGCAAGCAGAGAAATGTCGCC 695
Db 784 AspGlnAsnGluGlnGlyTyTyLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
QY 696 TCTGAACCAAAATCGTTACTGCTCTTACATCGTTTCTGGATGAGAAATGTCATTTTGA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
QY 756 CATACCAGTATATTACTGAGTTTCTTGTAGTGGCTGMAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyArgPheLeuGlnIle 843
QY 816 TATCTCTCAGAGACACAGCATAGAGTTCTCTCAATCGGAGACAACTATGAACTGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrgluLeuHis 863
QY 876 CTTTTCGACTACTCTTCAGAAACCTTCGGATCAGTATTGCTGCTTAAAGTGATA 932
Db 864 LeuLeuHisTyrgluGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882

RESULT 10
AAU74749
ID AAU74749 standard; protein; 882 AA.
AC AAU74749;
XX
DT 09-APR-2002 (first entry)
DE Human protease PRTS-9 protein sequence.
XX
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis.
XX
OS Homo sapiens.
XX
PN WO200198468-A2.
XX
PD 27-DEC-2001.
XX
PF 13-JUN-2001; 2001WO-US019178.
XX
PR 16-JUN-2000; 2000US-0212336P.
XX
PR 22-JUN-2000; 2000US-0213953P.
XX
PR 29-JUN-2000; 2000US-0215396P.
XX
PR 07-JUL-2000; 2000US-0216821P.
XX
PR 14-JUL-2000; 2000US-0218946P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
PI Deleageane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
PI Wallia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;
XX
PI WPI; 2002-090437/12.
XX
DR N-PSDB; ABK12892.
```

```
XX
PI Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
PI the diagnosis, treatment and prevention of gastrointestinal (e.g.
PI gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
PI (e.g. cancer) disorders.
PS Claim 1; Page 140-142; 177pp; English.
XX
CC The present invention relates to twenty one new human proteases, referred
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
CC invention are useful in the diagnosis, treatment and prevention of
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
CC disease and reproductive e.g. infertility and endometriosis disorders.
CC Numerous other examples of each disorder are given in the specification.
CC The present protein sequence represents the human protease PRTS-9 protein
CC of the invention
XX
SQ Sequence 882 AA;
XX
Alignment Scores:
Pred. No.: 3.1e-180 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: Gaps: 1
XX
US-10-825-632-4 (1-1197) x AAU74749 (1-882)
QY 3 TTGTAAGGCACCAAGACTCCCTTTAGAGCATCATCCTGTACGTAGTCAGTAAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrglyValValSerTyrglyValAsn 543
QY 63 CCTGGAGAGGTGACAGGCTGACTGACCTGGCTACTCACATCTTCTGTCAGTCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrglySerHisSerCysValSerGln 563
QY 123 CACTGTGACTCTTTTATAGTAAAGTATAGTAACAGAGAATCCACACTGTGTCTCCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrglySerAsnGlnLysAsnProHisCysValSerLeu 583
QY 183 TACAAGCTATCAAGTCTCTGAGATGACCACTTGCAAAACAAAGGAATTTTGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 243 ATTTTGGATTTCAGCAGGTCCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrglyProGluIlePheSerPheGlu 623
QY 303 AGTACTACTGGATTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 362
Db 624 SerThrThrGlyPheThrLeuTyrglyMetLeuTyrglyProHisAspLeuGlnProGly 643
QY 363 AGAAATATCTACTGCTGCTGCTATATATATGCTGCTCTCAG----- 404
Db 644 LysLysTyrglyProThrValLeuPheIleTyrglyGlyProGlnValGlnLeuValAsnAsn 663
QY 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrglyPheArgLeuAsnThrLeuAlaSerLeuGlyTyrglyVal 683
QY 404 ----- 404
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 405 -----GGTCAAAATAGAAATTCACGATCAGGTGGAGGACTCCAAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspGlnValGluGlyLeuGlyLeuGlnTyrglyLeuAlaSer 723
```





CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,  
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,  
CC relaxant and anticonvulsant. In addition, they are useful in screening  
CC assays to identify small molecules that modulate or inhibit, for example,  
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
CC used as in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein  
CC of the invention.  
XX  
SQ Sequence 882 AA;

Alignment Scores:  
Pred. No.: 3 1e-180 Length: 882  
Score: 1645.50 Matches: 310  
Percent Similarity: 86.4% Conservativity: 0  
Best Local Similarity: 86.4% Mismatches: 0  
Query Match: 76.9% Indels: 49  
DB: 5 Gaps: 1

US-10-825-632-4 (1-1197) x ADI117086 (1-882)

Qy 3 TTGAAGGACCAAGACTCCCTTAGAGCATCACCTGTACGTACTCAGTTACGTAAT 62  
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543  
Qy 63 CTGAGAGGTCAAGGCTGACGACCGTGGCTACTCACATCTCTGCTGCATCAGTAC 122  
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563  
Qy 123 CACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAATCCACACTGTGTGCCCTT 182  
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
Qy 183 TACAAGCTATCAAGTCTGGAAGATGACCCAACTTTGCAAAACAAAGAAATTTGGCCACC 242  
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
Qy 243 ATTTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 302  
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623  
Qy 303 AGTACTACTGGATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACAGCTGGA 362  
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
Qy 363 AAGAAATATCTACTGTGCTGTTCATATATGCTGTGCTCTCAG----- 404  
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663  
Qy 404 ----- 404  
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
Qy 404 ----- 404  
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
Qy 405 -----GGTCAATAGAAATTGACATCAGGTGGAGGACTCCAAATATCTAGCTTCT 455  
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
Qy 456 CGATATGATTTGATTCATAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGATAC 515  
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyIleTyrSerTyrGlyGlyTyr 743  
Qy 516 CTCTCCCTGATGCAATTAATGACAGGTCAGATATCTTCAGGGTGTCTATGCTGGGGCC 575  
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
Qy 576 CCAGTCACTGTGGATCTTCTATGATACAGGATACAGGAACCTTATATGGGTCACTT 635  
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783

Qy 636 GACCAGAAATGAACACAGGGCTATTACTTGGATCTCTGGCCATGCAACAGAAAGTTCCCC 695  
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803  
Qy 696 TCTGAACCAAAATCGTTTACTCTCTTACATGTTTCTGATGAGATGTCATTTTGA 755  
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluHisValHisPheAla 823  
Qy 756 CATACCAATATATTACTGAGTGTCTTGTAGTGGCTCGAAAGCCATATGATTTACAGATC 815  
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843  
Qy 816 TATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTAAGACTGCAT 875  
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863  
Qy 876 CTTTTGCACCTACCTTCAAGAAACCTTGGATCAGCTATTGCTCTTAAAGAGTATA 932  
Db 864 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882

RESULT 12

ABU07720  
ID ABU07720 standard; protein; 882 AA.  
XX

AC ABU07720;

XX 19-MAY-2003 (first entry)

XX Human serine protease HIPHUM46.

XX Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;  
KW serine protease activity modulation; dipeptidyl peptidase activity;  
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;  
KW Alzheimer's disease; parasupranuclear palsy; Huntington's disease;  
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;  
KW irritable bowel syndrome; type I diabetes; fascial incontinence; anaemia;  
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;  
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;  
KW multiple sclerosis.

XX Homo sapiens.

XX Key Location/Qualifiers  
XX 259..260 /note= "Paired glutamates of the beta propeller domain"

FT Region 739

FT Active-site /label= Catalytic\_serine\_residue

FT Active-site 817 /label= Catalytic\_aspartate\_residue

FT Active-site 849 /label= Catalytic\_histidine\_residue

XX GB2374869-A.

XX 30-OCT-2002.

XX 22-JAN-2002; 2002GB-00001404.

XX 23-JAN-2001; 2001GB-00001760.

XX (GLAX ) GLAXO GROUP LTD.

XX Edbrooke MR, Lewis AP;

XX WPI; 2003-150703/15.

XX N-PSDB; ABX12255.

XX Identifying modulators of serine protease activity useful for treating  
XX musculoskeletal diseases, by contacting cell expressing a novel serine  
XX protease polypeptide with a compound and monitoring serine protease  
XX activity.

PS Claim 10; Page 26-29; 38pp; English.

XX The invention relates to a method of identifying a substance that  
 CC modulates serine protease activity, comprising contacting a cell such as  
 CC a neuronal cell, lung cell, intestinal cell or a cell infected with a  
 CC virus, expressing a serine protease polypeptide (HIPHUM 46), or its  
 CC variant having dipeptidyl peptidase activity, or a serine protease  
 CC isolated from the cell with a test substance, and monitoring for serine  
 CC protease activity. The method is useful for identifying a substance that  
 CC modulates serine protease activity. A modulator of the serine protease is  
 CC useful in the manufacture of a medicament for treatment or prophylaxis of  
 CC a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus  
 CC infection, Alzheimer's disease, parasupranuclear palsy, myotonic  
 CC dystrophy, Huntington's disease or amyotrophic lateral sclerosis.  
 CC Additional disease that may be treated using modulators of the serine  
 CC protease include malabsorption syndromes, irritable bowel syndrome, lung  
 CC disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,  
 CC rectal polyps, small bowel tumours, colorectal tumours, anaemia,  
 CC dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple  
 CC sclerosis. The present sequence represents the amino acid sequence of the  
 CC human serine protease HIPHUM46  
 XX  
 SQ Sequence 882 AA;

Alignment Scores:  
 Pred. No.: 3.1e-180 Length: 882  
 Score: 1645.50 Matches: 310  
 Percent Similarity: 86.4% Conservative: 0  
 Best Local Similarity: 86.4% Mismatches: 0  
 Query Match: 76.9% Indels: 49  
 DB: 6 Gaps: 1

US-10-825-632-4 (1-1197) x ABU07720 (1-882)

QY 3 TTTRAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAGTAAT 62  
 DB 524 PHEGLUGLYTHRYLSEASPERPROLEUGLHSHISLEUTYRVALSERTYRVALASN 543  
 QY 63 CTGTGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATCTTGTCTGCATCAGTCAG 122  
 DB 544 PROGLYGLUVALTHRARGLEUTHRASPARGGLYTYRSERHISSERCYSYLIESERGLN 563  
 QY 123 CACTGTGACTCTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 182  
 DB 564 HISCYASPPHEPHEILESERLYSTYRSERASNGLNLYSASNPROHISCYSVASLSERLEU 593  
 QY 183 TACAACCTATCAAGTCTGACATGACCCCACTTGCAGAAACAAAGGAATTTTGGCCACC 242  
 DB 584 TYRLEUSERSESPROGLUASPPROTHRCYSLYSTHRLYSGLUPHETTPALATHR 603  
 QY 243 ATTTTGGATTACGACAGGTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302  
 DB 604 ILELEUASPSERALAGLYPROLEUPROASPTYRTHRPROGLIULIEPHESESPHEGLU 623  
 QY 303 AGTACTACTGGATTACATTGATGGGATGCTTACAGCCTCAGATCTACAGCCTGGA 362  
 DB 624 SERTHRTHRGLYPHETHRLEUTYRGLYMETLEUTYRYSPROHISASPLEUGLINPROGLY 643  
 QY 363 AAGAAATATCTACTGTGCTGTTCATATATATGCTGCTCAG----- 404  
 DB 644 LYSLEYTYRPROTHRVALHEUPHEILETYRGLYGLYPROGLINVALGLINLEUVALASN 663  
 QY 404 ----- 404  
 DB 664 ARGPHELYSGLYVALYSTYRPHARGLEUASNTHRLAUASERLEUGLYTYRVALVAL 693  
 QY 404 ----- 404  
 DB 684 VALVALILEASPSASNARGGLYSERCYSHISARGGLYLEULYSPHEGLUGLYALAPHELYS 703  
 QY 405 -----GGTCAATAGAAATTCAGCATCAGGTGAGGACTCCAAATATCTAGTTCT 455  
 DB 704 TYRLESMETGLYGLINLEUGLYLEASPSGLINVALGLUGLYLEUGINTYRLEUVALASER 723

QY 456 CGATATGATTTTCTTACCTAGTCTGTGGGCATCCAGCGTGGTCTCTATGGAGGATAC 515  
 DB 724 ARGTYRASPHEILEASPLEUASPARGVALGLYILEHISGLYTRIPSEPTYRGLYGLYTYR 743  
 QY 516 CTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGCC 575  
 DB 744 LEUSERLEUMETALALEUAMETGLNARGSERASPILEPHEARGVALALALEALAGLYALA 763  
 QY 576 CCAAGTCACTCTGTGGATCTTCATATGATACAGGATACAGGAACGTTATATGGGTACCCCT 635  
 DB 764 PROVALTHRLAUTRPILEPHEFYRASPETHRYTYRTHRGUARGTYRMETGLYHISPRO 783  
 QY 636 GACCAGATCAACAGGCGCTATTACTTACGATCTGTGGCCATCGACAGCAGAAAGTTCCCC 695  
 DB 784 ASPGLNASNGLUNGGLYTYRTYRLEUGLYSERVALALAMETGLNALAGLULYSPEPRO 803  
 QY 696 TCTGAACCAATCGTTTACTGTCTTACATGCTTCTTGGATGAGAAATGTCATTTCACA 755  
 DB 804 SERGLUPROASNARGLEULEULEUHLISGLYPHELEUASPLEUASNVALHISPEALALA 823  
 QY 756 CATACCAAGTATATTACTGAGTCTTTTGTAGTGGGCTGGAAAGCCATATGATTTACAGATC 815  
 DB 824 HISTHRSERILELEULEUSERPHELEUVALARGALAGLYLYSPROTYRASPLeUGLNILE 843  
 QY 816 TATCTCTCAGCAGACACACAGCATATAGAGTTCTCGAATCGGAGAACATTATGACTGCAT 875  
 DB 844 TYRPROGLINGLUARGHISERILEARGVALPROGLIUSERGLYGLUHISTYRGLULEU 863  
 QY 876 CTTTGTGCACTACTCTTCAAGAAACCTTGGATCACGTCATTGCTGTCTCTAAAGTGATA 932  
 DB 864 LEULEUHISTYRLEUGLINGLUASNLEUGLYSERARGLEALALEULEULYSVALILE 882

RESULT 13  
 ADIL17085  
 ID ADIL17085 standard; protein; 883 AA.  
 XX  
 AC ADIL17085;  
 DT  
 DT 15-APR-2004 (first entry)  
 XX  
 DE Murine NOVX protein homologue SeqID 621.  
 XX  
 KW mouse; murine; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;  
 KW inflammation; autoimmune disorder; allergy; blood disorder;  
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
 KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;  
 KW Alzheimer's disease; infect.  
 XX  
 OS Mus musculus.  
 XX  
 FN WO200268649-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 31-JAN-2002; 2002WO-US002785.  
 XX  
 PR 31-JAN-2001; 2001US-0265395P.  
 PR 31-JAN-2001; 2001US-0265412P.  
 PR 31-JAN-2001; 2001US-0265514P.  
 PR 31-JAN-2001; 2001US-0265517P.  
 PR 02-FEB-2001; 2001US-0266406P.  
 PR 05-FEB-2001; 2001US-0266767P.  
 PR 07-FEB-2001; 2001US-0266975P.  
 PR 07-FEB-2001; 2001US-0267057P.  
 PR 08-FEB-2001; 2001US-0267459P.  
 PR 08-FEB-2001; 2001US-0267823P.  
 PR 15-FEB-2001; 2001US-0268974P.  
 PR 26-FEB-2001; 2001US-0271664P.  
 PR 27-FEB-2001; 2001US-0271839P.  
 PR 27-FEB-2001; 2001US-0271855P.  
 PR 02-MAR-2001; 2001US-0272788P.  
 PR 02-MAR-2001; 2001US-0273046P.  
 PR 14-MAR-2001; 2001US-0275925P.

PR 14-MAR-2001; 2001US-0275947P.  
PR 14-MAR-2001; 2001US-0275950P.  
PR 14-MAR-2001; 2001US-0275989P.  
PR 15-MAR-2001; 2001US-0276448P.  
PR 15-MAR-2001; 2001US-0276450P.  
PR 16-MAR-2001; 2001US-0276397P.  
PR 16-MAR-2001; 2001US-0276768P.  
PR 20-MAR-2001; 2001US-0278652P.  
PR 26-MAR-2001; 2001US-0278752P.  
PR 26-MAR-2001; 2001US-0278778P.  
PR 26-MAR-2001; 2001US-0279882P.  
PR 29-MAR-2001; 2001US-0279884P.  
PR 30-MAR-2001; 2001US-0280147P.  
PR 11-APR-2001; 2001US-0282992P.  
PR 11-APR-2001; 2001US-0283083P.  
PR 20-APR-2001; 2001US-0285133P.  
PR 23-APR-2001; 2001US-0285749P.  
PR 03-MAY-2001; 2001US-0288327P.  
PR 03-MAY-2001; 2001US-0288504P.  
PR 29-MAY-2001; 2001US-0294047P.  
PR 30-MAY-2001; 2001US-0294473P.  
PR 08-JUN-2001; 2001US-0296964P.  
PR 18-JUN-2001; 2001US-0298959P.  
PR 19-JUN-2001; 2001US-0299324P.  
PR 13-AUG-2001; 2001US-0312020P.  
PR 16-AUG-2001; 2001US-0312889P.  
PR 16-AUG-2001; 2001US-0312908P.  
PR 21-AUG-2001; 2001US-0313390P.  
PR 28-AUG-2001; 2001US-0313470P.  
PR 31-AUG-2001; 2001US-031447P.  
PR 07-SEP-2001; 2001US-0318115P.  
PR 07-SEP-2001; 2001US-0318118P.  
PR 12-SEP-2001; 2001US-0318740P.  
PR 18-SEP-2001; 2001US-0323379P.  
PR 18-OCT-2001; 2001US-0330245P.  
PR 18-OCT-2001; 2001US-0330308P.  
PR 14-NOV-2001; 2001US-0332701P.  
XX PA (CURA-) CURAGEN CORP.  
XX  
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;  
PI Li L, Gangolli EA, Padigara M, Anderson DW, Rastelli L, Miller CE;  
PI Gerlach VL, Taupier RJ, Gusev VV, Colman SD, Wolenc AR, Pena CE;  
PI Furtak K, Groesse WM, Alsbrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX WPI; 2002-706998/76.  
XX  
XX New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
XX Disclosure; SEQ ID NO 621; 1498pp; English.  
XX  
XX This invention relates to a novel nucleic acids, and encoded polypeptides  
CC thereof, which have properties related to the stimulation of biochemical  
CC or physiological responses in a cell, tissue, organ or organism.  
CC Specifically, it refers to the use of biologically active fragments for  
CC diagnostic and prognostic assays and furthermore in the treatment of  
CC diverse pathological conditions. The present invention describes novel  
CC human and murine NOVX proteins, as well as methods to modulate their  
CC expression using antisense oligos, ribozymes and peptide nucleic acids.  
CC The NOVX polypeptides, polynucleotides and antibodies are useful in  
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,  
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in  
CC treating or preventing diseases such as inflammation, autoimmune  
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome  
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,  
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy  
CC and epilepsy. Accordingly, these molecules have many activities including  
CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,  
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,  
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic, 695  
CC  
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,  
CC relaxant and anticonvulsant. In addition, they are useful in screening  
CC assays to identify small molecules that modulate or inhibit, for example,  
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
CC used as in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein  
XX of the invention.  
SQ Sequence 883 AA;  
Alignment Scores:  
Pred. No.: 1,31e-175 Length: 883  
Score: 1605.50 Matches: 302  
Percent Similarity: 85.0% Conservative: 3  
Best Local Similarity: 84.1% Mismatches: 5  
Query Match: 75.0% Indels: 49  
DB: 5 Gaps: 1  
US-10-825-632-4 (1-1197) x ADI17085 (1-883)  
Qy 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTTACGTAAAT 62  
Db 525 PheGluGlyThrIysAspSerProLeuGluHisIleuTyrValThrSerTyrAlaAsn 544  
Qy 63 CCTGGAGAGGTGACAGGCTGACTGACCGTGGTACTCACATTTCTGCTGCATCAGTCAG 122  
Db 545 ProGlyGluValValArgLeuThrAspArgGlyTyrSerHisSerCysCysLeuSerArg 564  
Qy 123 CACTGTGACTTCTTTAAGTAAGTATAGTAACGAGAGAGATCCACATCTGTGTCTCTT 182  
Db 565 HisCysAspPhePheIleSerIysTyrSerAsnGlnLysAsnProHisCysValSerLeu 584  
Qy 183 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACC 242  
Db 585 TyrLysLeuSerSerProGluAspProValHisLysThrLysGluPheTyrPalaThr 604  
Qy 243 ATTTTGGATTTCAGCAGGTCCTCTCTGACTATATATGTTGCTGCTCTCAG----- 404  
Db 605 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 624  
Qy 303 AGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGAGCCTCATGATCTACAGCCTGGA 362  
Db 625 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 644  
Qy 363 AAGAAATATCTACTGCTGCTGTTTATATATGTTGCTGCTCTCAG----- 404  
Db 645 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 664  
Qy 404 ----- 404  
Db 665 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 684  
Qy 404 ----- 404  
Db 685 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 704  
Qy 405 -----GGTCAATAGAAATTGACCATCAGGTGGAGGACTCCATATCTAGCTTCT 455  
Db 705 TyrLysMetGlyGlnIleGluIleAspAsnGlnValGluGlyLeuGlnTyrLeuAlaSer 724  
Qy 456 CGATATGATTTCACTTACATTCGTGTGGGCATCCACGGCTGCTGCTATGAGAGGATAC 515  
Db 725 GlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyrGlyTyr 744  
Qy 516 CTCTCCCTGATGGCAATTAATCAGAGGTCAGATATCTTACAGGTTGCTATGCTGGGGCC 575  
Db 745 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 764  
Qy 576 CCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAGGAAACGTTATATGGGTCAACCT 635  
Db 765 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 784  
Qy 636 GACCAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATCTGAGCAAGCAAGAAAGTTCCCC 695

Db 785 AppGlnAsnGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 804  
Qy 696 TCTGAACCAATCGTTTACTGCTTACATGCTTCTCGGATGAGAAATGTCATTTTGCA 755  
Db 805 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 824  
Qy 756 CATACAGATATATTACTGAGTTTATAGTGGGCTGGAAGCCATATGATTACAGATC 815  
Db 825 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 844  
Qy 816 TATCTCAGAGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTATGACTGCAT 875  
Db 845 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 864  
Qy 876 CTTTTGCACTACCTTCAAGAAAACCTTGGATCACGATTTGCTGCTCTAAAAGTGATA 932  
Db 865 LeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 883  
RESULT 14  
ADE78977  
ID ADE78977 standard; protein; 493 AA.  
XX  
AC ADE78977;  
DT 29-JAN-2004 (first entry)  
XX  
DE Human protein modification and maintenance molecule (PMMW) -15.  
XX  
KW protein modification and maintenance molecule; PMMW;  
KW protein modification; protein maintenance; protein function;  
KW protein conformation; protein stabilisation; protein degradation; kinase;  
KW phosphatase; protease; protease inhibitor; isomerase; transferase;  
KW molecular chaperone; anti-HIV; anti-allergic; anti-inflammatory;  
KW antianaemic; antiparkinsonian; nootropic; anticonvulsant;  
KW antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;  
KW cytotropic; hepatotropic; dermatological; antidiabetic; nephrotropic;  
KW antigout; thymimetic; neuroprotective; osteopathic; antiarthritic;  
KW antiparasitic; antihelminthic; antipsoriatic; uropathic; ophthalmological;  
KW antirheumatic; haemostatic; antibacterial; virucide; protozoacide;  
KW fungicide; gene therapy; cell proliferative disorder; arteriosclerosis;  
KW hepatitis; polycythaemia vera; psoriasis; primary thrombocytopaenia;  
KW cancer; developmental disorder; anaemia; mental retardation;  
KW neurological disorder; Alzheimer's disease; Parkinson's disease;  
KW epilepsy; autoimmune disorder; inflammatory disorder; AIDS; allergies;  
KW asthma; autoimmune thyroiditis; Crohn's disease; diabetes mellitus;  
KW glomerulonephritis; Goodpasture's syndrome; multiple sclerosis;  
KW arthritis; osteoporosis; pancreatitis; Sjogren's syndrome;  
KW microbial infection; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2003063688-A2.  
XX  
PD 07-AUG-2003.  
XX  
PF 23-JAN-2003; 2003WO-US0002500.  
XX  
PR 25-JAN-2002; 2002US-0351928P.  
PR 25-FEB-2002; 2002US-0359903P.  
PR 21-MAR-2002; 2002US-0366837P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Hafalia AJA, Li JX, Gorvad AE, Chawla NK, Sprague WW, Lee SY;  
PI Chang H, Elliott VS, Ramkumar J, Khare R, Emerling EM, Kable AE;  
PI Tang YT, Yue H, Gietzen KO, Lee S, Swarnakar A, Baughn MR;  
PI Wilson AD, Jin P, Chien D, Hawkins PR, Jiang X, Jackson AA;  
PI Bhatia U, Burrill JD, Blake JJ, Ho A, Zheng W, Ison CH, Marquis JP;  
PI Tran UK, Lal PG, Warren BA, Xu Y, Honchell CD, Becha SD;  
PI Lehr-Mason PM;  
XX  
DR WPI; 2003-636761/60.

DR N-PSDB; ADE79035.  
XX  
PT New human protein modification and maintenance molecules and  
PT polynucleotides, useful for diagnosing, treating or preventing autoimmune  
PT or inflammatory disorders (e.g. AIDS, allergy or anemia), multiple  
PT sclerosis or cancer.  
XX  
PS Claim 1; SEQ ID NO 15; 405pp; English;  
XX  
CC This invention relates to novel isolated human proteins, which are human  
CC protein modification and maintenance molecules (PMMW). The cellular  
CC processes regulating modification and maintenance of protein molecules  
CC coordinate their function, conformation, stabilisation and degradation.  
CC Each of these processes is mediated by key enzymes or proteins such as  
CC kinases, phosphatases, proteases, protease inhibitors, isomerases,  
CC transferases and molecular chaperones. Compounds which modulate the  
CC proteins of the invention may have anti-HIV, anti-allergic,  
CC antiinflammatory, antianaemic, antiparkinsonian, nootropic,  
CC anticonvulsant, antiarteriosclerotic, antiasthmatic, immunosuppressive,  
CC antithyroid, cytotropic, hepatotropic, dermatological, antidiabetic,  
CC nephrotropic, antigout, thymimetic, neuroprotective, osteopathic,  
CC antiarthritic, antiparasitic, antihelminthic, antipsoriatic, uropathic,  
CC ophthalmological, antirheumatic, haemostatic, antibacterial, virucide,  
CC protozoacide or fungicide activities. The DNA sequence which encodes the  
CC proteins of the invention may be useful for gene therapy. The human  
CC protein modification and maintenance molecules (PMMWs), the DNA sequences  
CC which encode them and their modulating compounds are useful for  
CC diagnosing, treating or preventing disorders associated with aberrant  
CC expression of PMMW, particularly cell-proliferative disorders (for  
CC example arteriosclerosis, hepatitis, polycythaemia vera, psoriasis,  
CC primary thrombocytopaenia or cancer), developmental disorders (for  
CC example anaemia or mental retardation), neurological disorders (for  
CC example Alzheimer's disease, Parkinson's disease or epilepsy),  
CC autoimmune/inflammatory disorders (for example AIDS, allergies, asthma,  
CC autoimmune thyroiditis, Crohn's disease, diabetes mellitus,  
CC glomerulonephritis, Goodpasture's syndrome, multiple sclerosis,  
CC arthritis, osteoporosis, pancreatitis, Sjogren's syndrome) or microbial  
CC infections. The present sequence is the amino acid sequence of a human  
CC PMMW of the invention.  
XX  
SQ Sequence 493 AA;  
Alignment Scores:  
Pred. No.: 8.92e-153 Length: 493  
Score: 1407.00 Matches: 270  
Percent Similarity: 88.4% Conservative: 11  
Best Local Similarity: 84.9% Mismatches: 19  
Query Match: 65.7% Indels: 18  
DB: 7 Gaps: 5  
US-10-825-632-4 (1-1197) x ADE78977 (1-493)  
Qy 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATACCTGCTAGCTAGTACGTTACCTAAT 62  
Db 186 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 205  
Qy 63 CCTGGAGAGGTGACAAAGGCTGACTGACCTGGCTACTCACATTTCTGTGTCATGAGT 122  
Db 206 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 225  
Qy 123 CACTGTGACTTCTTTATAGTAAAGTATAGTAACCGAAGAATCCACACTGTGTGTCCTTT 182  
Db 226 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 245  
Qy 183 TACAGCTATCAGTCTCTGAAGTATGACCACTCCAAACAAAGAAATTTTGGCCACC 242  
Db 246 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheThrPalatnr 265  
Qy 243 ATTTTGGATTACAGAGTCTCTTCTCTGACTACTCTCTCCAGAAATTTTCTCTTTGAA 302  
Db 266 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 285  
Qy 303 AGTACTACTGGATTTACATTGTTATGGATGCTCTACAGCCTCATGATCTACAGCCTGA 362

Db 286 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 305  
Qy 363 AGAAATATCTACTGTGCTGTTCATATATATGGTCTCTCAGGGTCAAAATAGAAATGAC 422  
Db 306 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu---ValAsn 324  
Qy 423 GATCAGGTGGAAGGCTCCATATAT-----CTAGCTTCT---CGATATGAT 464  
Db 325 AsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal 344  
Qy 465 TTCATTGACTTGTAGCTCGTGGGCATC-----CACGGCTGGTCTATGGAGATACCTC 518  
Db 345 ValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPhe 364  
Qy 519 TCCTGATGGCATTAAATCAGAGGTGCAGATATCTTCAGGGTTGCTATTGTGGGGCCCA 578  
Db 365 LysTyrLys-----MetValAlaIleAlaGlyAlaPro 375  
Qy 579 GTCACTCTGTGATCTTCTATGATACAGGATACACGGAAGTATTATGGTCACTGAC 638  
Db 376 ValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAsp 395  
Qy 639 CAGATGACAGGGCTATTACTTAGGATCTGTGGCCATGCACGAGAAAGTCCCTCT 598  
Db 396 GlnAsnGlnGlnGlyTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSer 415  
Qy 699 GAACCAATCGTTTACTGCTCTTACATGGTTCTCGATGAGAAATGTCATTTGCAAT 758  
Db 416 GluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHis 435  
Qy 759 ACAGTATATTACTAGTCTTTTATGAGGCTCGAAGCCATATGATTTACAGATCTAT 818  
Db 436 ThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyr 455  
Qy 819 CCTCAGGAGACACAGCATAAGATTCCTGAATCGGGAGAACATATTGAACTGCATCTT 878  
Db 456 ProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeu 475  
Qy 879 TTGCATCTCTTCAAGAAACCTTGGATCAGTATTGCTGCTTAAAGTGATA 932  
Db 476 LeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 493

## RESULT 15

ABU92026  
ID ABU92026 standard; protein; 831 AA.

AC ABU92026;

XX ABU92026;

XX 15-JUL-2003 (first entry)

XX 15-JUL-2003 (first entry)

XX Human protein modification and maintenance molecule-6 (PMMW-6).

XX Human; protein modification and maintenance molecule; PMMW; cancer;  
XX cell proliferation disorder; atherosclerosis; neurological disorder;  
XX epilepsy; Huntington's disease; stroke; immune disorder; allergy;  
XX inflammatory disorder; AIDS; developmental disorder; hypothyroidism;  
XX Cushing's syndrome; gastrointestinal disorder; epithelial disorder;  
XX infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
XX neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;  
XX antiinflammatory; thyromimetic.

XX Homo sapiens.

XX OS

XX WO2003031939-A2.

XX PN

XX 17-APR-2003.

XX PD

XX 11-OCT-2002; 2002WO-US032850.

XX PF

XX 12-OCT-2001; 2001US-0329689P.

XX PR

XX 25-OCT-2001; 2001US-0335703P.

XX PR

XX 09-NOV-2001; 2001US-0348887P.

PR 28-NOV-2001; 2001US-0334145P.

PR 06-DEC-2001; 2001US-0337451P.

PR 14-DEC-2001; 2001US-0340584P.

XX (INCY-) INCYTE GENOMICS INC.

XX Rankumar J, Gorvad AB, Baughn MR, Emerling BM, Yang J, Lee SY;

XX Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;

XX Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;

XX Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;

XX Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;

XX WPI; 2003-430274/40.

XX N-PSDB; ACA92421.

XX New human protein modification and maintenance molecules (PMMW), useful

XX for diagnosing, treating and preventing diseases or conditions associated

XX with the aberrant PMMW expression e.g. cancer, atherosclerosis, or

XX infections.

XX Claim 1; Page 242-243; 311pp; English.

XX The present invention relates to the isolation of human protein

XX modification and maintenance molecules (PMMW) and the polynucleotide

XX sequences encoding them. A total of 40 PMMW polypeptides (designated PMMW

XX 1 to PMMW-40) are disclosed. The sequences of the invention are useful

XX for diagnosing a condition or disease associated with the expression of

XX PMMW in a subject, preparing a polyclonal or monoclonal antibody, and

XX generating an expression profile of a sample containing the

XX polynucleotides. The diseases or conditions associated with decreased

XX expression or overexpression of PMMW are cell proliferation disorders

XX (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,

XX Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,

XX allergies), developmental disorders (e.g. hypothyroidism, Cushing's

XX syndrome), gastrointestinal or epithelial disorders, and infections. The

XX PMMW polypeptides or their fragments are useful in screening compounds

XX for effectiveness as agonists or antagonists of the polypeptides, or in

XX altering the expression of the target polynucleotide and compounds that

XX specifically bind to, or modulate the activity of the polypeptide.

XX ABU92021-ABU92060 represent the human PMMW polypeptides of the invention

XX SQ Sequence 831 AA;

## Alignment Scores:

Pred. No.: 1.16e-152 Length: 831  
Score: 1407.00 Matches: 270  
Percent Similarity: 88.4% Conservative: 11  
Best Local Similarity: 84.9% Mismatches: 19  
Query Match: 65.7% Indels: 18  
DB: Gaps: 5

US-10-825-632-4 (1-1197) x ABU92026 (1-831)

Qy 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCAGCTAGTACGTAGTACGTAGTAAAT 62

Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543

Qy 63 CCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCAGATTTCTGCTGATCAGTCAG 122

Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563

Qy 123 CACTGTGACTTCTTTATAGTATAGTAAACCGAAGAATCCACACTGTGTGTCCTT 182

Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583

Qy 183 TACAAGCTATCAAGTCCTGAAGATGACCCAACTTGCAGAAACAAAGGAATTTGGCCACC 242

Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603

Qy 243 ATTTTGGATTACAGAGGTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA 302

Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623

```
QY 303 A G T A C T A C T G G A T T T A C A T T G T A T G G G A T C C T C T C A A G C C T C A T G A T C A C G C C T G G A 362
Db |||||||
QY 624 S e r T h r T h r G l y P h e T h r L e u T y r G l y M e t L e u T y r L y s P r o H i s A s p L e u G l n P r o G l y 643
Db |||||||
QY 363 A A G A A T A T C C T A C T G T G C T G T C A T A T A T G G T G G T C C T C A G G G T C A A A T A G A A T T G A C 422
Db |||||||
QY 644 L y s L y s T y r P r o T h r V a l L e u P h e I l e T y r G l y P r o G l n V a l G l n L e u --- V a l A s n 662
Db |||||||
QY 423 G A T C A G G T G G A A G A C T C C A A T A T --- C T A G C T T C T --- C G A T A T G A T 464
Db |||||||
QY 663 A s n A r g P h e L y s G l y V a l L y s T y r P h e A r g L e u A s n T h r L e u A l a S e r L e u G l y T y r V a l 682
Db |||||||
QY 465 T T C A T T G A C T T A G A T C G T G T G G G A T C --- C A C G G T G G T C C T A T G A G G A G A T A C C T C 518
Db |||||||
QY 683 V a l V a l V a l I l e A s p A s n A r g G l y S e r C y s H i s A r g G l y L e u L y s P h e G l u G l y A l a P h e 702
Db |||||||
QY 519 T C C C T G A T G G C A T T A A T G C A G A G G T C A G A T A T C T C A G G G T T G C T A T T G C T G G G C C C C A 578
Db |||||||
QY 703 L y s T y r L y s --- M e t V a l A l a I l e A l a G l y A l a P r o 713
Db |||||||
QY 579 G T C A C T C T G T G G A T C T T C T A T G A T A C A G G A C G T T A T A T G G G T C A C C C T G A C 638
Db |||||||
QY 714 V a l T h r L e u T r p I l e P h e T y r A s p T h r G l y T y r T h r G l u A r g T y r M e t G l y H i s P r o A s p 733
Db |||||||
QY 639 C A G A A T G A A C A G G G C T A T T A C T T A G G A T C T G T G G C C A T C C A A G C A G A A A A G T T C C C T C T 698
Db |||||||
QY 734 G l n A s n G l u G l n G l y T y r T y r L e u G l y S e r V a l A l a M e t G l n A l a G l u L y s P h e P r o S e r 753
Db |||||||
QY 699 G A A C C A A T C G T T T A C T G C T T T A C A T G G T T T C T G G A T G A G A A T G C C A T T T T G C A C A T 758
Db |||||||
QY 754 G l u P r o A s n A r g L e u L e u L e u H i s G l y P h e L e u A s p G l u A s n V a l H i s P h e A l a H i s 773
Db |||||||
QY 759 A C C A G T A T A T T A C T G A G T T T T T A G T C A G G G C T G G A A G C C A T A T A T G A T T T A C A G A T C T A T 818
Db |||||||
QY 774 T h r S e r I l l e L e u L e u S e r P h e L e u V a l A r g A l a G l y L y s P r o T y r A s p L e u G l n I l e T y r 793
Db |||||||
QY 819 C C T C A G G A G A C A C A C A T A A G A G T T C C T G A T C G G A G A A C A T T A T G A A C T G C A T C T T 878
Db |||||||
QY 794 P r o G l n G l u A r g H i s S e r I l l e A r g V a l P r o G l u S e r G l y G l u H i s T y r G l u L e u H i s L e u 813
Db |||||||
QY 879 T T G C A C T A C C T T C A A G A A A C C T T G G A T C A C G T A T T G C T G C T C T A A A A G T G A T A 932
Db |||||||
QY 814 L e u H i s T y r L e u G l n G l u A s n L e u G l y S e r A r g I l e A l a A l a L e u L y s V a l I l e 831
Db |||||||
```

Search completed: May 2, 2006, 01:03:09  
Job time : 200.717 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 02:50:19 ; Search time 4.00111 Seconds  
(without alignments)  
2721.092 Million cell updates/sec

Title: US-10-825-632-4  
Perfect score: 2140  
Sequence: 1 attttgaaggaccacaaagac.....ttactaaaaaaaaaaaaa 1197

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 464238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB\_spool/US10825632/runat\_01052006\_110000\_3533/app\_query.fasta.1  
-DB=Published Applications\_AA\_New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -HOST=abes02p  
-USER=US10825632 @CGN\_1\_1\_59 @runat\_01052006\_110000\_3533 -NCPU=6 -ICPU=3  
-NO\_MMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_AA\_New.\*

1: /SIDSS5/ptodata/1/pubpaa/US08\_NEW\_PUB.pep1.\*  
2: /SIDSS5/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /SIDSS5/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /SIDSS5/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
5: /SIDSS5/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
6: /SIDSS5/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
7: /SIDSS5/ptodata/1/pubpaa/US09\_NEW\_PUB.pep1.\*  
8: /SIDSS5/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
9: /SIDSS5/ptodata/1/pubpaa/US10\_NEW\_PUB.pep1.\*  
10: /SIDSS5/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
11: /SIDSS5/ptodata/1/pubpaa/US11\_NEW\_PUB.pep1.\*  
12: /SIDSS5/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1645.5	76.9	882	11	US-11-151-601-20
2	1093.5	51.1	497	11	US-11-264-096-1594
3	1089.5	50.9	497	11	US-11-264-096-1593
4	577	27.0	109	11	US-11-176-951-10
5	427.5	20.0	745	11	US-11-079-463-6408
					Sequence 20, Appl
					Sequence 1594, Ap
					Sequence 1593, Ap
					Sequence 10, Appl
					Sequence 6408, Ap

6	394	18.4	99	11	US-11-176-951-11	Sequence 11, Appl
7	341.5	16.0	760	11	US-11-208-288-6	Sequence 6, Appli
8	341	15.9	627	11	US-11-079-463-7758	Sequence 7758, Ap
9	337.5	15.8	738	11	US-11-208-288-4	Sequence 4, Appli
10	337.5	15.8	766	9	US-10-501-035-234	Sequence 234, App
11	337.5	15.8	766	11	US-11-208-288-2	Sequence 2, Appli
12	333.5	15.6	762	11	US-11-116-939-13	Sequence 13, Appli
13	333.5	15.6	766	9	US-10-522-789-2	Sequence 2, Appli
14	302	14.1	760	11	US-11-186-284-55	Sequence 55, Appl
15	180.5	8.4	624	11	US-11-079-463-7504	Sequence 7504, Ap
16	164	7.7	349	11	US-11-264-096-1591	Sequence 1591, Ap
17	152.5	7.1	657	11	US-11-179-977-1	Sequence 1, Appli
18	145	6.8	102	11	US-11-176-951-7	Sequence 7, Appli
19	142.5	6.7	737	11	US-11-079-463-9281	Sequence 9281, Ap
20	139	6.5	115	11	US-11-176-951-9	Sequence 9, Appli
21	139	6.5	115	11	US-11-176-951-12	Sequence 12, Appli
22	133.5	6.2	102	11	US-11-176-951-8	Sequence 8, Appli
23	106.5	5.0	315	11	US-11-018-868-143	Sequence 143, App
24	105.5	4.9	352	11	US-11-098-686-11006	Sequence 11006, A
25	98.5	4.6	316	11	US-11-079-463-9481	Sequence 9481, Ap
26	98.5	4.6	668	9	US-10-454-437-118	Sequence 118, App
27	95	4.4	314	11	US-11-179-977-17	Sequence 17, Appl
28	93.5	4.4	256	11	US-11-179-977-2	Sequence 2, Appli
29	92	4.3	397	11	US-11-084-508-5	Sequence 5, Appli
30	91	4.3	507	11	US-11-188-298-5830	Sequence 5830, Ap
31	88.5	4.1	255	11	US-11-179-977-3	Sequence 3, Appli
32	88	4.1	342	11	US-11-087-099-5518	Sequence 5518, Ap
33	88	4.1	1147	9	US-10-453-372-4	Sequence 4, Appli
34	88	4.1	1179	9	US-10-912-580-1	Sequence 1, Appli
35	88	4.1	1181	9	US-10-912-580-2	Sequence 2, Appli
36	88	4.1	1181	9	US-10-912-580-3	Sequence 3, Appli
37	87	4.1	308	11	US-11-094-568A-33777	Sequence 33777, A
38	84.5	3.9	309	11	US-11-094-568A-32738	Sequence 32738, A
39	84.5	3.9	335	11	US-11-094-568A-32737	Sequence 32737, A
40	83.5	3.9	344	11	US-11-079-463-7255	Sequence 7255, Ap
41	83.5	3.9	420	11	US-11-079-463-6314	Sequence 6314, Ap
42	83.5	3.9	2323	9	US-10-793-626-760	Sequence 760, App
43	83.5	3.9	3130	9	US-10-453-372-42	Sequence 42, Appl
44	83.5	3.9	3483	9	US-10-453-372-40	Sequence 40, Appl
45	83.5	3.9	3546	9	US-10-453-372-32	Sequence 32, Appl

#### ALIGNMENTS

RESULT 1  
US-11-151-601-20  
; Sequence 20, Application US/11151601  
; Publication No. US20060003413A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Welch, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.  
; FILE REFERENCE: MF100-054P/KCP/OMIDVIM  
; CURRENT APPLICATION NUMBER: US/11/151,601  
; CURRENT FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: US 10/170,789  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06529  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/882,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15

```
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Alignment Scores:
Pred. No.: 1,648-160 Length: 882
Score: 1645.50 Matches: 310
Percent Similarity: 86.4% Conservative: 0
Best Local Similarity: 86.4% Mismatches: 0
Query Match: 76.9% Indels: 49
DB: 11 Gaps: 1

US-10-825-632-4 (1-1197) x US-11-151-601-20 (1-882)

Qy 3 TTGAAGGACCAAGATCCCTTTAGAGCATCACCTGTACGTAGTCAGTTCAGTAAAT 62
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy 63 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGCTCATCAGTCAG 122
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysHisLeuSerGln 563
Qy 123 CACTGTGACTCTTTTAAAGTAAATAGTAACTAACAGAGAAATCCACACTGTGTGCTCTT 182
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 183 TACAAGCTATCAAGTCTCAAGATCACCACTTGCAAAACAAAGGAATTTTGGGCCACC 242
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTpaIaThr 603
Qy 243 ATTTTGGATTTCAGAGTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 302
Db 604 IleLeuaspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
Qy 303 AGTACTACTGGATTACATTGTGGATGTCTTACAGCTCATGATCTACAGCCTGGA 362
Db 624 SerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 363 AGAAATATCTACTGTGCTGCTTCATATATATGCTGCTCTCAG----- 404
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 404 ----- 404
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 404 ----- 404
Db 684 ValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 405 -----GGTCAATAGAAATATGACGATCAGGTGGAAGACTCCCAATATCTAGCTTCT 455
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 456 CGATATGATTTCAATTCATTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATAC 515
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPsrTyrGlyTyr 743
Qy 516 CTCTCCCTGATGCATTAATGAGAGGTTCAGATATCTTCAGGGTGTGCTATTGTGGGGCC 575
Db
```

```
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 576 CCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATGGTTCACCCCT 635
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
Qy 636 GACCAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATCTGCAAGCAGAAAAGTTCCCC 695
Db 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
Qy 696 TCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAAATGTCCTATTTGCA 755
Db 804 SerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAla 823
Qy 756 CATACCAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATC 815
Db 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
Qy 816 TATCTCTCAGGAGACACAGCATTAAGCTTCTGAATCGGAGAACATTTATGAATGCAT 875
Db 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
Qy 876 CTTTTCGACTACTCTTCAGAAACCTTCGATCAGTATTCGCTCTAAAAGTGATA 932
Db 864 LeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIleAlaLeuLysValIle 882

RESULT 2
US-11-264-096-1594
; Sequence 1594, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; PRIOR FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1594
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-264-096-1594

Alignment Scores:
Pred. No.: 6,77e-104 Length: 497
Score: 1093.50 Matches: 200
Percent Similarity: 69.3% Conservative: 42
Best Local Similarity: 57.3% Mismatches: 58
Query Match: 51.1% Indels: 49
DB: 11 Gaps: 1

US-10-825-632-4 (1-1197) x US-11-264-096-1594 (1-497)

Qy 3 TTGAAGGACCAAGATCCCTTTAGAGCATCACCTGTAGTACGTAGTTCAGTTCAGTAAAT 62
Db 149 PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAla 169
Qy 63 CTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAG 122
Db 169 AlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGln 189
Qy 123 CACTGTGACTCTTTTAAAGTAAATAGTAACTAACAGAAAGATCCACACTGTGTGCTCCTT 182
Db 189 AsnPheAspMetPheValSerHisTyrSerValSerThrProProCysValHisVal 208
```

```
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCACAAACAAAGAAATTTTGGCCACC 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAlaSer 228
QY 243 ATTTGGATTACAGAGTCTCTCTCTGCTACTATCTCTCCAGAAATTTTCTTTTGA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 MetMetGluAlaAlaSerCysProProAspTyrValProGluIlePheHisPheHis 248
QY 303 AGTACTACTGGATTACATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 362
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 268
QY 363 AAGAAATATCTACTGCTGCTCATATATGCTGCTCTCTCTCTCTCTCTCTCTCTCT 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269 LysLysHisProThrValLeuPheValTyrGlyProGlnValGlnLeuValAsnAsn 288
QY 404 ----- 404
Db 289 SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 308
QY 404 ----- 404
Db 309 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 328
QY 405 -----GGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCAATATCTAGCTTCT 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 348
QY 456 CDATATGATTTCATGATTAGATTCGTGTGGGCATCCAGCGTGTCTATGAGGATAC 515
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyPhe 368
QY 516 CTCTCCCTGATGCAATATGACAGGTGATATCTTCTCAGGTTGCTATGCTGGGGCC 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 388
QY 576 CAGTCACTCTGTGGGATCTTCATGATACAGGATACAGGACGTTATATGGTCAACCT 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 408
QY 636 GACCAGAAATCAACAGCGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTCTCCC 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 GluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 428
QY 696 TCTGAACCAATCGTTTACTGCTTACATGTTTCTCGATGATGAAATGTCATTTGCA 755
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 429 AsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePhe 448
QY 756 CATACCAATATATCTAGTATTTTATGAGGGCTGGAAGCCATATGATTTACAGATC 815
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 468
QY 816 TATCTCAGAGAGACACAGCATAGAGTTCTCTGATCGGAGACATATGAACTGCAT 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 TyrProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThr 488
QY 876 CTTTTGCACTACCTTCAAGAAACCTT 902
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 489 LeuLeuHisPheLeuGlnGluTyrLeu 497
```

## RESULT 3

```
US-11-264-096-1593
; Sequence 1593, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1593
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (183)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-264-096-1593
```

## Alignment Scores:

Pred. No.:	1.74e-103	Length:	497
Score:	1089.50	Matches:	199
Percent Similarity:	69.1%	Conservative:	42
Best Local Similarity:	57.0%	Mismatches:	59
Query Match:	50.9%	Indels:	49
DB:	11	Gaps:	1

US-10-825-632-4 (1-1197) x US-11-264-096-1593 (1-497)

```
QY 3 TTTGAAGGCCCAAGACCTCCCTTTAGAGCATCACTGTAGTAGTCAGTTACGTAAT 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValIleSerTyrGluAla 168
QY 63 CTGCGAGGTGACACAGGCTGACTGACCGTGGCTACTCACATTTGTCGATCAGTCAG 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 AlaGlyGluIleValArgLeuThrThrProGlyPheSerHis***CysSerMetSerGln 188
QY 123 CACTGTGACTTCTTTTAAAGTAAGTAGTAACACAGAGAAATCCACACTGTGTGCCCT 182
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 AsnPheAspMetPheValSerHisTyrSerValSerValSerThrProCysValHisVal 208
QY 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCACAAACAAAGAAATTTTGGCCACC 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 TyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAlaSer 228
QY 243 ATTTGGATTACAGAGTCTCTCTCTGATATCTCTCCAGAAATTTTCTTTTGA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 MetMetGluAlaAlaSerCysProProAspTyrValProGluIlePheHisPheHis 248
QY 303 AGTACTACTGGATTACATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 362
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 268
QY 363 AAGAAATATCTACTGCTGCTTCATATATGTTGGTGGTCTCTCAG----- 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269 LysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsn 288
QY 404 ----- 404
Db 289 SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 308
QY 404 ----- 404
Db 309 ValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLys 328
QY 405 -----GGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCAATATCTAGCTTCT 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 348
QY 456 CDATATGATTTCATGATTAGATTCGTGTGGGCATCCAGCGTGTCTATGAGGATAC 515
    :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyPhe 368
QY 516 CTCTCCCTGATGCAATATGACAGGTGATATCTTCTCAGGTTGCTATGCTGGGGCC 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
||||| 369 LeuSerLeuMetGlyLeuIleHisGlyProGlnValPheLysValAlaIleAlaGlyAla 388
||||| 576 CAGTCACTCTGATCTCTATGATACAGGATACACGGAAGTGTATATGGTCAACCT 635
||||| 389 ProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValPro 408
||||| 636 GACCAAGTAAACAGGGCTATTACTTAGGATCTCTGGCCATGCAAGCAGAAAAGTTCCCC 695
||||| 409 GluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuPro 428
||||| 696 TCTGAACCAAACTGTTTACTGCTTACATGTTTCTGATGAGAAATGCCATTTTGA 755
||||| 429 AsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePhe 448
||||| 756 CATACAGTATATTACTGAGTTTCTAGTGAGGCTGGAAGCCATATGATTTACAGATC 815
||||| 449 HisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIle 468
||||| 816 TATCCTCAGGAGACAGACAGCATAGAGTTCTGCAATCGGAGACATATTGACTGCAT 875
||||| 469 TyrProAsnGluArgHisSerIleArgCysProGluSerGlyGluHisTyrGluValThr 488
||||| 876 CTTTGCATCTACTCTCAAGAAACCTT 902
||||| 489 LeuLeuHisPheLeuGlnGluTyrLeu 497
```

## RESULT 4

```
US-11-176-951-10
; Sequence 10, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: PROLYL DIPEPTIDASES
; CURRENT APPLICATION NUMBER: US/11/176,951
; PRIOR FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-10
```

```
Alignment Scores:
Pred. No.: 5,128-51 Length: 109
Score: 577.00 Matches: 109
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 27.0% Indels: 0
DB: 11 Gaps: 0
```

US-10-825-632-4 (1-1197) x US-11-176-951-10 (1-109)

```
Qy 606 GGATACCGGAACGTTATATGGTCCCTGACCAAGTAAAGCAAGGCTATTACTAGGA 665
Db 1 GlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGly 20
Qy 666 TCTGTGGCCATCAACGCAAGAAAGTTCCCTCTCTGAAACCAATCGTTTACTGCTTACAT 725
Db 21 SerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHis 40
Qy 726 GGTTCCTCGATGAGAATGTCATTTTGCATACACCATGATATATTACTGAGTTTTTATGTG 785
Db 41 GlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuVal 60
```

```
Qy 786 AGGCTGGAAGCAACCATATGATTACAGATCTATCTCAGGAGACACAGCATTAAGATT 845
Db 61 ArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgVal 80
Qy 846 CCTGAATCGGAGAAACATTATGAACCTGCATCTTTTGCACTACCTTCAAGAAAACCTTGA 905
Db 81 ProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGly 100
Qy 906 TCACGTATTGCTGCTCTAAAGTGATA 932
Db 101 SerArgIleAlaAlaLeuLysValIle 109
RESULT 5
US-11-079-463-6408
; Sequence 6408, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Brelton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRAC
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6408
; LENGTH: 745
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-6408
```

```
Alignment Scores:
Pred. No.: 1,76e-35 Length: 745
Score: 427.50 Matches: 97
Percent Similarity: 46.8% Conservative: 66
Best Local Similarity: 27.9% Mismatches: 124
Query Match: 20.0% Indels: 61
DB: 11 Gaps: 6
```

US-10-825-632-4 (1-1197) x US-11-079-463-6408 (1-745)

```
Qy 3 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTACGTACGTAAAT 62
Db 411 TyrThrSerAsnGluGluSerProMetArgGlnAlaValTyrLysIleAsp-----Arg 428
Qy 63 CTGGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTCACATTCTTGTGTCAGTCAG 122
Db 429 LysGlyLysLysMetLysLeuSerAsnGlnProGlyThrAsnSerProIlePheSerSer 448
Qy 123 CACTGTGACTCTTTTATAGTATAGTAACCAAGAAATCCACACTGTGTGTGCTCCTT 182
Db 449 SerMetLysTyrPheMetAsnLysPheThrSerLeuAspThrProMetLeuIleThrLeu 468
Qy 183 TACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAAGGAATTTTGGGCCACC 242
Db 469 -----AsnAspAsnThrGlyLysValLeuLysThrLeuValThr 481
Qy 243 ATTTTGATTACAGAGGTCTCTTCTGACTATATCTCTCCA-----GAAATTTTCTCT 296
Db 482 AsnAspLysLeuLysGlnLysLeuAlaGluTyrAlaIleProGlnLysGluPhePheThr 501
Qy 297 TTTGAAAGTACTACTCGGATTTACATTGTATGGATGCTCTACAGCCCTCATGATCTACAG 356
Db 502 PheLysThrThrGluGlyValAspLeuAsnGlyTyrMetLysProValAsnPheAsp 521
Qy 357 CTGGAAAGAAATATCTACTGTGTGTTTCATATATGTTGCTCTCAGGGTCAATAGAA 416
Db 522 ProAlaLysArgTyrProValLeuMetPheGlnTyrSerGlyProGlySerGlnGlnVal 541
Qy 417 ATTGAC----- 422
```



Db 486 ArgValLeuGlu-----AspIleSerAlaLeuLeuArgMetLeuGln 499  
QY 270 GACTATACCTCCAGAA-----ATTTCTCTTTTGAAGTACTACTGGATTAC 320  
Db 500 AspValGlnMetProSerLysLysLeuAspPheIleValLeuAsnGluThrArgPhe--- 518  
QY 321 TTGTATGGGATCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTG 380  
Db 519 TrpTyrGlnMetIleLeuProPheHis---PheAspLysSerLysLysTyrProLeuLeu 537  
QY 381 CTGTTTCATATATGGTGGCTCT-----PheAspLysSerLysLysTyrProLeuLeu 401  
Db 538 LeuAspValTyrAlaGlyProCysSerGlnLysAlaAspAlaSerPheArgLeuAsnTrp 557  
QY 401 ----- 401  
Db 558 AlaThrTyrLeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySer 577  
QY 402 -----CAGGGTCAAAATAGAAAT 419  
Db 578 GlyTyrGlnGlyAspLysIleMetHisAlaIleAsnArgArgLeuGlyThrLeuGluVal 597  
QY 420 GAGCATCAGTGAAGGACTCCATATCTAGCTTCTCGATATGATTTTCTTGGACTTAGAT 479  
Db 598 GluAspGlnIleGluAlaArgGlnPheVal---LysMetGlyPheValAspSerLys 616  
QY 480 CTGTGGGATCCAGCGTGTCTTATGAGGATACCTCTCCCTGATGGCATTAAATGAC 539  
Db 617 ArgValAlaIleTrpGlyTyrSerTyrGlyTyrValThrSerMetValLeuGlySer 636  
QY 540 AGGTGATATCTTCCAGGTGTCTTCTCTGGGCGCCAGTCACTCTGGGATCTTCTAT 599  
Db 637 GlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGluTyrTr 656  
QY 600 GATACAGGATACACGAGCGTTATATGGT-----CACCTGACCAAGTAATGAACAGGC 653  
Db 657 AspSerValTyrThrGluArgTyrMetGlyLeuProIleProGluAsnLeuLeuAspHis 676  
QY 654 TATTACTTAGTCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 713  
Db 677 TyrArgAsnSerThrValMetSerArgAlaGluHisPhe-----LysGlnValGluTyr 694  
QY 714 CTGCTCTTACATGGTTCTCTGGATGAGAATGTCCATTTTGCACATACCACTATATTACTG 773  
Db 695 LeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGlnGlnSerAlaGlnIleSer 714  
QY 774 AGTTTCTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC 833  
Db 715 LysAlaLeuValAspAlaGlyValAspPheGlnAlaMetTrpTyrThrAspGluAspHis 734  
QY 834 AGCATAAAGATTCTCGAATCGGAGAACATTATCACTGCTTTTGGCACTACCTTCAA 893  
Db 735 GlyIleAlaSerSerThrAlaHisGlnHisIleTyrSerHisMetSerHisPheLeuGln 754  
QY 894 GAA 896  
Db 755 Gln 755

## RESULT 8

US-11-079-463-7758  
; Sequence 7758 Application US/11079463  
; Publication No. US20060073161A1  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR  
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: PATH00-03DIV2  
; CURRENT APPLICATION NUMBER: US/11/079,463  
; CURRENT FILING DATE: 2005-03-14  
; PRIOR FILING DATE: 60/128,705  
; PRIOR FILING DATE: 1999-04-09  
; PRIOR APPLICATION NUMBER: US 09/540,209  
; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 10444  
; SEQ ID NO 7758  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: B.fragilis  
US-11-079-463-7758  
Alignment Scores:  
Pred. No.: 1,28e-26 Length: 627  
Score: 341.00 Matches: 92  
Percent Similarity: 43.2% Conservative: 60  
Best Local Similarity: 26.1% Mismatches: 126  
Query Match: 15.9% Indels: 74  
DB: 11 Gaps: 9

US-10-825-632-4 (1-1197) x US-11-079-463-7758 (1-627)

QY 12 ACCAAAGACTCCCTTTAGAGCATCACCTGTAGTACTAGTAAATCTCTGGAGAG 71  
Db 295 ThrGluPheSerProLeuGlnAsnAsnLeuPheArgLeuAsp---ThrLysThrGlyThr 313  
QY 72 GTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTCATCAGTCAGCAGCTGTGAC 131  
Db 314 ArgThrProLeuGlySerAlaGluGlyValHisSerGlyGlnLeuSerProSerGlyArg 333  
QY 132 TTCTTTTAAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTA 191  
Db 334 TyrLeuIleAspGlnTyrAsnSerProThrValProArgSerIleAsnIleIleAspVal 353  
QY 192 TCAAGTCTCGAAGATGACCACTTGGCAAAACAAAGGAATTTTGGGCCACCATTTTGGAT 251  
Db 354 GlnSerGlyLysSer-----ValAsnLeuLeuThr 363  
QY 252 TCAGCAGGTCCTCTTCTGACTATATCTCTCTCA-----GAAATTTCTCTTTTGAAGT 305  
Db 364 AlaAlaAspProPheThrGlyTyrLysMetProGlyIleGluThrGlyThrIleLysAla 383  
QY 306 ACTACTGGATTACA---TTGATGGGATGCTCTAAGACCTCATGATCTACAGCTCGGA 362  
Db 384 AlaAspGlyLysThrAspLeuTyrArgLeuIleLysProAlaAspPheAspProAsn 403  
QY 363 AAGAAATATCTACTGCTGTGTATATATGTTGCTCTCTCAGGCTCAATA----- 413  
Db 404 LysLysTyrProAlaIleValTyrValTyrGlyGlyProHisAlaGlnLeuValThrAsn 423  
QY 413 ----- 413  
Db 424 GlyTrpGlnAsnGlyAlaArgGlyTrpAspIleTyrMetAlaAsnLysGlyTyrIleMet 443  
QY 413 ----- 413  
Db 444 PheThrValAspGlyArgGlySerSerAspArgGlyLeuAspPheGluAsnValThrPhe 463  
QY 414 -----GAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCT 455  
Db 464 ArgGlnLeuGlyIleGluGluGlyArgAspGlnValLysGlyThrGluPheLeuLysSer 483  
QY 456 CGATATGATTCTATTCAGCTTAGATCGTGTGGGCATCCACGGCTGCTCTCTATGGAGGATAC 515  
Db 484 ---LeuProTyrValAspGlyAsnArgIleGlyValHisGlyTrpSerPheGlyGlyHis 502  
QY 516 CTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCTCAGGTTGCTATTGCTGGGGCC 575  
Db 503 MetThrAlaLeuLeuLeuArgTyrProGluIlePheLysValGlyValAlaGlyGly 522  
QY 576 CAGTCACTCTGTGGATCTTCTATATACAGGATACGGAACGTATATGGTCAACCT 635  
Db 523 ProValIleAspTrpGlyTyrTrpGluValMetTyrGlyGluArgTyrMetAspThrPro 542  
QY 636 GACCAGAAATGAACAGGGCTTATTACTTAGATCTGTGGCATCGAAGCAGAAAAGTTTCCCC 695  
Db 543 GlnSerAsnProLysGlyTyrLysGluCysAsnLeuLeuLysAsnLeuAlaGlyAsnLeu---- 561





```

Query Match: 15.8% Indels: 75
DB: 9 Gaps: 13
US-10-825-632-4 (1-1197) x US-10-501-035-234 (1-766)

QY 15 AAAGACTCCCTTTAGAGCATCACCTGCTACGTACGTACGTAAATCTCGAGAGGTG 74
DB 423 LysGlyMetProGlyGlyArgAsnLeuThrLysLeuLeuLeuLeuLeuVal 442
QY 75 ACAAGGCTGACTGACCGTGGCTACTCATTCTGCTGCTGCTGCTGCTGCTGCTG 131
DB 443 ThrCysLeu-----SerCysGluLeuAsnProGluArgCysGln 455
QY 132 TTTCTTTAAGTAAGTATAGTAACACGAGAATCCACAC-----TGT----- 173
DB 456 TyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGlyPro 475
QY 174 ---GTGTCCTTTTACAGCTATCAAGTCTGAGAGTACCACTTGCACCAAGGAA 230
DB 476 GlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArgValLeuGlu 495
QY 231 TTTTGGGCCACCATTTGGATTTCAGCAGGTCTCTCTGACTATCTCTCCAGAA--- 287
DB 496 Asp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLys 512
QY 288 -----ATTTTCTCTTTGAAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAAG 341
DB 513 LysLeuAspPheLeileLeuAsnGluThrLysPhe---TrpTyrGlnMetLeuPro 531
QY 342 CCTCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTCATATATGTTGCTCT 401
DB 532 ProHis---PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyPro 550
QY 401 ----- 401
DB 551 CysSerGlnLysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThr 570
QY 401 ----- 401
DB 571 GluAsnLeileValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIle 590
QY 402 -----CAGGTCCTCAATAGAAATTTGACGATCAGTGGGCTCCAGGACTC 440
DB 591 MethisAlaIleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAla 610
QY 441 CAATATCTAGCTTCTCGATATGATTTTCTGATCTTAGATCGTGGGCTCCAGGCTGG 500
DB 611 ArgGlnPhe---SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrp 629
QY 501 TCTATGGAGGATACCTCTCCCTGATGCTATTAATGCAGAGGTGAGATATCTTCAGGTT 560
DB 630 SerTyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCys 649
QY 561 GCTATTGCTGGGCCCCAGTCTCTGTGGATCTTCTATGATACAGGATACAGGAACTG 620
DB 650 GlyIleAlaValAlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArg 669
QY 621 TATATGGGT-----CACCTGACCGATGAGAGGCTATTACTAGGATCTGTGGCC 674
DB 670 TyrMetGlyLeuProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMet 689
QY 675 ATGCACGAGAAAGTTCCCTCTGAAACCAATCGTTTACCTCTTACATGTTCTTCTG 734
DB 690 SerArgAlaGluAsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAla 707
QY 735 GATGAGAAATGCTCAATTTTGACATACCACTATATCTAGTGTGTTTATGAGGCTGGA 794
DB 708 AspAspValHisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGly 727
QY 795 AAGCCATATGATTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCCTGAATCG 854
DB 728 ValAspPheGlnAlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAla 747

```

```

QY 855 GGAGAACATTATGAACATCATCTTTTGCACCTACCTTCAAGAA 896
DB 748 HisGlnHisIleTyrThrHisMetSerHisPheIleLysGln 761
RESULT 11
US-11-208-288-2
; Sequence 2, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; PRIOR FILING DATE: 2005-08-18
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-2
Alignment Scores:
Pred. No.: 3,07e-26 Length: 766
Score: 337.50 Matches: 99
Percent Similarity: 42.9% Conservative: 53
Best Local Similarity: 28.0% Mismatches: 127
Query Match: 15.8% Indels: 75
DB: 11 Gaps: 13
US-10-825-632-4 (1-1197) x US-11-208-288-2 (1-766)
QY 15 AAAGACTCCCTTTAGAGCATCACCTGCTACGTACGTACGTAAATCTCGAGAGGTG 74
DB 423 LysGlyMetProGlyGlyArgAsnLeuThrLysLeuLeuLeuLeuVal 442
QY 75 ACAAGGCTGACTGACCGTGGCTACTCATTCTGCTGCTGCTGCTGCTGCTGCTG 131
DB 443 ThrCysLeu-----SerCysGluLeuAsnProGluArgCysGln 455
QY 132 TTTCTTTAAGTAAGTATAGTAACACGAGAATCCACAC-----TGT----- 173
DB 456 TyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGlyPro 475
QY 174 ---GTGTCCTTTTACAGCTATCAAGTCTGAGAGTACCACTTGCACCAAGGAA 230
DB 476 GlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArgValLeuGlu 495
QY 231 TTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTGACTATCTCTCCAGAA--- 287
DB 496 Asp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLys 512
QY 288 -----ATTTTCTCTTTGAAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAAG 341
DB 513 LysLeuAspPheLeileLeuAsnGluThrLysPhe---TrpTyrGlnMetLeuPro 531
QY 342 CCTCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTCATATATGTTGCTCT 401
DB 532 ProHis---PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyPro 550
QY 401 ----- 401
DB 551 CysSerGlnLysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThr 570
QY 401 ----- 401
DB 571 GluAsnLeileValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIle 590
QY 402 -----CAGGTCCTCAATAGAAATTTGACGATCAGTGGGCTCCAGGACTC 440

```

Db 591 MetHisAlaIleAsnArgLeuGlyThrPheGluValGluAspGlnIleGluAlaA 610  
Qy 441 CAATATCTAGCTTCTCGATATGATTTCAATGTAGTCTGTGGGCAATCCACGGCTGG 500  
Db 611 ArgGlnPhe---SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrp 629  
Qy 501 TGCTATGGAGGATACCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTCAGGGTT 560  
Db 630 SerTyGlyGlyTyValThrSerMetValLeuGlySerGlySerGlyValPheLysCys 649  
Qy 561 GCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATCACCGAACGT 620  
Db 650 GlyIleAlaValAlaProValSerArgTrpGluTyTyAspSerValTyThrGluArg 669  
Qy 621 TATATGGGT-----CACCTGCAGCAATGACAGGCTATTACTTAGGATCTGTGGCC 674  
Db 670 TyrMetGlyLeuProThrProGluAspAsnLeuAspHisTyArgAsnSerThrValMet 689  
Qy 675 ATGCAAGCAGAAAGTTCCTCTGACCAATCGTTTACTGCTTACATGTTCTCTG 734  
Db 690 SerArgAlaGluAsnPhe-----LysGlnValGluTyLeuLeuIleHisGlyThrAla 707  
Qy 735 GATGAGAATCTCATTTTGCACATACCACTATATTACTGAGTCTTTTGTAGTGAGGCTGGA 794  
Db 708 AspAspAsnValHisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGly 727  
Qy 795 AAGCCATATGATTTACAGATCTATCTCTCAGGAGACACACATAGAGTTCTCGAATCG 854  
Db 728 ValAspPheGlnAlaMetTrpTyThrAspGluAspHisGlyIleAlaSerSerThrAla 747  
Qy 855 GAGAACATATGAACTGCTCTTTGCACTACCTTCAAGAA 896  
Db 748 HisGlnHisIleTyThrHisMetSerHisPheIleLysGln 761

## RESULT 12

; Sequence 13, Application US/11116939  
; Publication No. US20050265995A1  
; GENERAL INFORMATION:  
; APPLICANT: Stephen Tomlinson  
; APPLICANT: Richard J. Quigg  
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS  
; FILE REFERENCE: 19113.011502  
; CURRENT APPLICATION NUMBER: US/11/116,939  
; PRIOR FILING DATE: 2005-04-28  
; PRIOR APPLICATION NUMBER: 60/565,907  
; PRIOR FILING DATE: 2004-04-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 762  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic  
; OTHER INFORMATION: construct  
US-11-116-939-13

Alignment Scores:  
Pred. No.: 7,89e-26 Length: 762  
Score: 333.50 Matches: 99  
Percent Similarity: 42.2% Conservative: 52  
Best Local Similarity: 27.7% Mismatches: 124  
Query Match: 15.6% Indels: 83  
DB: 11 Gaps: 14

US-10-825-632-4 (1-1197) x US-11-116-939-13 (1-762)

Qy 15 AAAGACTCCCTTTAGAGCATCCCTGTACGTAGTCTAGCTTAACGTAATCCTGAGAGGTG 74  
Db 419 LysGlyMetProGlyGlyArgAsnLeuTyLysIle----- 430  
Qy 75 ACAAGCTGACTGACCGTGGCTACTCACATCTTGTGTCATCAGT----- 119

## RESULT 13

US-10-522-789-2  
; Sequence 2, Application US/10522789  
; Publication No. US20050260732A1  
; GENERAL INFORMATION:  
; APPLICANT: TANABE SEIYAKU CO., LTD.

Db 431 ---GlnLeuSerAsp-----TyrThrLysValThrCysLeuSerCysGlnLeuAsnPro 447  
Qy 120 CAGCACTGTGACTCTTTTATAGTAAGTATAGTACACAGAAATCCACAC----- 170  
Db 448 GluArgCysGlnTyTySerValSerPheSerLysGluAlaLysTyTyGlnLeuArg 467  
Qy 171 TGT-----GTGTCCCTTTTCAAGCTATCAAGTCTCTGGAAGATGACCACTTGC 218  
Db 468 CysSerGlyProGlyLeuProLeuTyThrLeuHisSerSerValAsnAspLysGlyLeu 487  
Qy 219 AAAACAAGAAATTTTGGGCGACCATTTTGGATTCTCAGCAGGCTCTCTCTCCTACTACT 278  
Db 488 ArgValLeuGluAsp---AenSerAlaLeuAspLysMet-----LeuGlnAsnValGln 504  
Qy 279 CCTCCAGAA-----ATTTCCTCTTTTGAAGATCTACTTGGATTTCATTTGATGGG 329  
Db 505 MetProSerLysLysLeuAspPheIleIleLeuAsnGluThrLysPhe---TrpTyGln 523  
Qy 330 ATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTCTGCTGTGTCATATA 389  
Db 524 MetIleLeuProHis---PheAspLysSerLysLysTyProLeuLeuAspVal 542  
Qy 390 TATCGTGTCTCT----- 401  
Db 543 TyrAlaGlyProCysSerGlnLysAlaAspThrValPheArgLeuAsnTrpAlaThrTyr 562  
Qy 401 ----- 401  
Db 563 LeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyGln 582  
Qy 402 -----CAGGTCAAATAGAAATTCAGCATCAG 428  
Db 583 GlyAspLysIleMetHisAlaIleAsnArgArgLeuGlyThrPheGluValGluAspGln 602  
Qy 429 GTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTTGACTAGATCGTGGGC 488  
Db 603 IleGluAlaAlaArgGlnPhe---SerLysMetGlyPheValAspAsnLysArgIleAla 621  
Qy 489 ATCCACGGCTGTCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTTCAGAT 548  
Db 622 IleTrpGlyTrpSerTyGlyTyValThrSerMetValLeuGlySerGlySerGly 641  
Qy 549 ATCTTCAGGTTTCTATGCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGA 608  
Db 642 ValPheLysCysGlyIleAlaValAlaProValSerArgTrpGluTyTyAspSerVal 661  
Qy 609 TACACGGAACGTTATATGGT-----CACCTGCAGCAATGACACGGGCTATTACTTA 662  
Db 662 TyrThrGluArgTyMetGlyLeuProThrProGluAspAsnLeuAspHisTyArgAsn 681  
Qy 663 GGATCTGTGGCCATGCAAGCAGAAAGTTCCTCTCTGAACCAATTCGTTTACTGCTCTTA 722  
Db 682 SerThrValMetSerArgAlaGluAsnPhe-----LysGlnValGluTyLeuLeuIle 699  
Qy 723 CATGGTTTCTGATGAGAAATGTCATTTTGCACATACAGCATATTAATTAATGATTTT 782  
Db 700 HisGlyThrAlaAspAspAsnValHisPheGlnGlnSerAlaGlnIleSerLysAlaLeu 719  
Qy 783 GTGAGGCTCGAAAGCCATATGATTTACATCTCTCAGGAGACACAGCATAGA 842  
Db 720 ValAspValGlyValAspPheGlnAlaMetTrpTyThrAspGluAspHisGlyIleAla 739  
Qy 843 GTTCCTGAATCGGAGAACATATATGAACCTGTCATCTTTTGCATCTACCTTCAAGAA 896  
Db 740 SerSerThrAlaHisGlnHisIleTyThrHisMetSerHisPheIleLysGln 757

```
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-522-789-2

Alignment Scores:
Pred. No.: 7,9e-26 Length: 766
Score: 333.50 Matches: 99
Percent Similarity: 42.2% Conservative: 52
Best Local Similarity: 27.7% Mismatches: 124
Query Match: 15.6% Indels: 83
DB: 9 Gaps: 14

US-10-825-632-4 (1-1197) x US-10-522-789-2 (1-766)
QY 15 AAAGACTCCCTTTAGAGCATCACCTGTACGTACGTACGTTACGTTAAATCCTGGAGAGTG 74
Db 423 LysGlyMetProGlyGlyArgAsnLeuTyrLysIle :::: 434
QY 75 ACAAGGCTGACTGACCGTGGCTACTCACATCTTGTGTCATCAGT----- 119
Db 435 ---GlnLeuSerAsp-----TyrThrLysValThrCysLeuSerCysGluLeuAenPro 451
QY 120 CAGCACTGACTCTTTATAGTANGATATAGTATACAGAGATCCACAC----- 170
Db 452 GluArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyrGlnLeuArg 471
QY 171 TGT-----GTGTCCTTTAACAAGCTATCAAGTCTCAAGTCTGAGATGACCCCACTTGC 218
Db 472 CysSerGlyProGlyLeuProLeuTyrThrLeuHisSerSerValAenAspLysGlyLeu 491
QY 219 AAACAAGAAATTTGGGCCACCATTTTGGATTTCAGACAGGTCTCTCTCGACTACTAT 278
Db 492 ArGValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAenValGln 508
QY 279 CCTCAGAA-----ATTTCTCTCTTTGAAAGTACTACTGGAATTTACATTTGATGGG 329
Db 509 MetProSerLysLysLeuAspPheIleLeuAenGluThrLysPhe---TrpTyrGln 527
QY 330 ATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTGCTGTTCATA 389
Db 528 MetIleLeuProHis---PheAspLysSerLysLysTyrProLeuLeuLeuAspVal 546
QY 390 TATGGTGGTCCCT----- 401
Db 547 TyrAlaGlyProCysSerGlnLysAlaAspThrValPheArgLeuAenTrpAlaThrTyr 566
QY 401 ----- 401
Db 567 LeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGln 586
QY 402 -----CAGGTCANATAGNAATGACGATCG 428
Db 587 GlyAspLysIleMetHisAlaIleAsnArgLeuGlyThrPheGluValGluAspGln 606
QY 429 GTGAAGAGACTCCAATATCTAGCTCTCTCGATATGATTTTCATTTAGCTTAGATCGTGTGGC 488
Db 607 IleGluAlaArgGlnPhe---SerLysMetGlyPheValAspAenLysArgIleAla 625
QY 489 ATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAAGGTCAGAT 548
Db 626 IleTrpGlyTrpSerTyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGly 645
QY 549 ATCTTCAGGGTTGCTATTGCTGGGGCCCCAGCTACTCTGTGGATCTTCTATGATACAGGA 608
```

```
Db 646 ValPheLysCysGlyIleAlaValAlaProValSerArgTyrGluTyrTyrAspSerVal 665
QY 609 TACACGGAACGTTATATGGT-----CACCTGACCAAGATGAACAGGGCTATTACTTA 662
Db 666 TyrThrGluArgTyrMetGlyLeuProThrProGluAspAenLeuAspHisTyrArgAsn 685
QY 663 GGATCTGTGCCATGCAACGACAGAAAGTCCCTCTGAACCAAAATCGTTTACTGCTCTTA 722
Db 686 SerThrValMetSerArgAlaGluAenPhe-----LysGlnValGluTyrLeuLeuIle 703
QY 723 CATGGTTTCTGGATGAGATGTCATTTTGGACATACACAGTATATATTACTGAGTTTTTTA 782
Db 704 HisGlyThrAlaAspAenValHisPheGlnGlnSerAlaGlnIleSerLysAlaLeu 723
QY 783 GTCAGGCTGGAAGCCATATGATTTACAGATCTCTCTCAGGAGACACAGCATAGA 842
Db 724 ValAspValGlyValAspPheGlnAlaMetTrpTyrThrAspGluAspHisGlyIleAla 743
QY 843 GTTCCTGAATCGGGAGAACATTATGAACCTGTCATCTTTTGCACACTACCTTCAAGAA 896
Db 744 SerSerThrAlaHisGlnHisIleTyrThrHisMetSerHisPheIleLysGln 761

RESULT 14
US-11-186-284-55
; Sequence 55, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schleser, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgeat, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-029PZRNW
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-186-284-55

Alignment Scores:
Pred. No.: 1,35e-22 Length: 760
Score: 302.00 Matches: 93
Percent Similarity: 40.5% Conservative: 55
Best Local Similarity: 25.5% Mismatches: 139
Query Match: 14.1% Indels: 78
DB: 11 Gaps: 11

US-10-825-632-4 (1-1197) x US-11-186-284-55 (1-760)
QY 3 TTGAAGGACCAAGAGCTCCCTTTAGAGCATACCTGTGATAGTCAGTTCAGTTCAGTAAT 62
Db 401 PheArgValThrGlnAspSerLeuPheTyrSerSerAenGluPheGluGluTyrProGly 420
QY 63 CCTGGAGAGGTGACAAAGGCTGACTGACCGCTGCTACTCACATTCT---TGCTGCATCAGT 119
```

```
Db      421  ArgArgAsnIleTyrArgIleSerIleGlySerTyrProSerLysLysCysValThr 440
QY      120  CAGCAC-----TGTCACCTCTTTATAGTAAGTATAGTAACTAACACAGAGAT 164
Db      441  CysHisLeuArgLysGluArgCysGlnTyrThrAlaSerPheSerAspTyrAlaLys 460
QY      165  CCACACTGTGTCTCCCTTTAC-----AAGCTATCAAGTCTCTGAAGATGACCCA 212
Db      461  TyrTyrAlaLeuValCysTyrGlyProGlyIleProIleSerThrLeuHisAspGlyArg 480
QY      213  ACTGCAAAACAAGAAATTTGGGCCACCAATTTTGATTCACAGAGTCTCTCTCTGAC 272
Db      481  ThrAspGlnGlu-----IleLysIleLeuGluGluAsnLysGluLeuGluAsn 496
QY      273  -----TATACTCTCCAGAAATTTCTCTTTTGAAAGTACTACTGGATT 317
Db      497  AlaLeuLysAsnIleGlnLeuProLysGluGluIleLysLysLeuGluValAspGluIle 516
QY      318  ACATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACT 377
Db      517  ThrLeuTyrTyrLysMetIleLeuProGlnPheAspArgSerLysLysTyrProLeu 536
QY      378  GTGCTGTTCAATATGTTGGTCT-----GTCCT----- 401
Db      537  LeuIleGlnValTyrGlyProCysSerGlnSerValArgSerValPheAlaValAsn 556
QY      401  ----- 401
Db      557  TrpIleSerTyrLeuAlaSerLysGluGlyMetValIleAlaLeuValAspGlyArgGly 576
QY      402  -----CAGGTCAAATAGAA 416
Db      577  ThrAlaPheGlnGlyAspLysLeuLeuTyrAlaValTyrArgLysLeuGlyValTyrGlu 596
QY      417  ATGACGATCAGGTGAAGGACTCCAATATCTAGCTCTCGATATGATTTTCATGACTTA 476
Db      597  ValGluAspGlnIleThrAlaValArgLysPheIle---GluMetGlyPheIleAspGlu 615
QY      477  GATCGTGTGGGATCCAGCGTGTCTATCGAGATACCTCTCCCTGATGGCATTAATG 536
Db      616  LysArgIleAlaIleTyrTrpSerTyrGlyTyrValSerSerLeuAlaLeuAla 635
QY      537  CAGAGTGCAGATATCTCAGGTTGCTATTGCTGGGCCCCAGTCACCTCTGTGGATCTTC 596
Db      636  SerGlyThrGlyLeuPheLysCysGlyIleAlaValAlaProValSerSerTrpGluTyr 655
QY      597  TATGATACAGGATACACGGAAGCTTATATGGTGCACCT-----GACCAAGATGAACAG 650
Db      656  TyrAlaSerValTyrThrGluArgPheMetGlyLeuProThrLysAspAspAsnLeuGlu 675
QY      651  GdCTATTACTTAGGATCTGTGGCCATGCAACGACGAAAGTTCCCTCTGAACCAATCGT 710
Db      676  HisTyrLysAsnSerThrValMetAlaArgAlaGluTyrPheArgAsnValAsp----- 693
QY      711  TTACTCTCTTACATGTTCTCTGGATGAGATGTCAATTTTGACATACACGATATATTA 770
Db      694  TyrLeuLeuIleHisGlyThrAlaAspAsnValHisPheGlnAsnSerAlaGlnIle 713
QY      771  CTGAGTTTTTATGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGA 830
Db      714  AlaLysAlaLeuValAsnAlaGlnValAspPheGlnAlaMetTrpTyrSerAspGlnAsn 733
QY      831  CACAGCATAGAGTCTCTGAATCGGGA-----GAACATTATGAAGTCTCTTTTG 881
Db      734  HisGlyLeu-----SerGlyLeuSerThrAsnHisLeuTyrThrHisMetThr 749
QY      882  CACTACTCTCAAGAA 896
Db      750  HisPheLeuLysGln 754
```

RESULT 15

US-11-079-463-7504

```
; Sequence 7504, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gaty L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRAT
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7504
; LENGTH: 624
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-7504
Alignment Scores:
Pred. No.: 3,85e-10 Length: 624
Score: 180.50 Matches: 57
Percent Similarity: 38.5% Conservative: 35
Best Local Similarity: 23.8% Mismatches: 84
Query Match: 8.4% Indels: 63
DB: 11 Gaps: 6
```

US-10-825-632-4 (1-1197) x US-11-079-463-7504 (1-624)

```
QY      291  TTCTCTTTTGAAGTACTACTCGATTTCATTTGATGGATGCTCTACACGCTCATGAT 350
Db      356  TrpAsnPheThrAlaSerAspGlyThrGluIleLysGlyMetCysLeuProSer 375
QY      351  CTACACCTCGGAAGAAATATCTACTGTCGTGTTTCATATATGGTGGT----- 398
Db      376  PheAspProAsnLysLysTyrProLeuIleValTyrTyrGlyThrProThr 395
QY      398  ----- 398
Db      396  GluArgGlyIleSerAsnProTyrCysAlaGlnLeuPheAlaSerArgAspTyrValVal 415
QY      399  -----CCTCAGGTCNAATA----- 413
Db      416  TyrValIleGlnProSerGlyThrIleGlyPheGlyGlnGluPheSerAlaArgHisVal 435
QY      414  -----GAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 455
Db      436  AsnAlaTrpGlyLysArgThrAlaAspAspIleIleGluGlyThrLysGlnPheCysLys 455
QY      456  CGATATCATTTCACTTAGACTTCGTGGGCATCCACGGCTGCTCTATGGAGGATAC 515
Db      456  GluHisProPheValAsnAspLysLysIleGlyCysLeuGlyAlaSerTyrGlyGlyPhe 475
QY      516  CTCTCCCTGATGCATTATGACAGGTCAGATATCTTCAGGGTGTCTATGCTGGGGCC 575
Db      476  MetThr---GlnTyrLeuGlnThrGlnThrAspIlePheAlaAlaValSerHisAla 494
QY      576  CCAGTCACCTCTGGGATCTTCTATGATACAGGATACAGGAACGTTATATGGT----- 629
Db      495  GlyIleSerAsnValThrSerTyr-----TrpGlyGluGlyTyrTrpGlyTyrGly 511
QY      630  -----CACCTGACCAAGATGAACAGGGCTATTACTTACTTA 662
Db      512  TyrAsnAlaIleAlaAlaAspSerTyrProTyrAsnAsnProGluPheThrLys 531
QY      663  GGATCTGTGGCCATGACAGCAGAAAGTTCCTCTGACCAACAAATCGTTTACTGCTCTTA 722
Db      532  GlnGlySerLeuPheAsnAlaAspLysIleAsnThrPro-----LeuLeuLeuLeu 548
QY      723  CATGGTTTCTCGATGAGATGCCATTTTGCACATACCATATATTTACTGAGTTTTCTTA 782
Db      549  HisGlyThrValAspThrAsnValProIleGlyGluSerIleGlnLeuPheAsnAlaLeu 568
```

Qy 783 GTGAGGCTGGAAGCCATATGATTACAGATCTATCCTCAGGAGAGACACAGCATA 839  
Db 569 LyeIleLeuGlyLysThrValGluPheIleThrValAspGlyGluAsnHisPheIle 587

Search completed: May 2, 2006, 03:10:38  
Job time : 37.0056 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:53:29 ; Search time 8318.8 Seconds  
(without alignments)  
11404.501 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 1669  
Sequence: 1 aacagtgacagcaatccta.....aaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: gb\_ba.\*
- 2: gb\_in.\*
- 3: gb\_env.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pr.\*
- 9: gb\_ro.\*
- 10: gb\_sts.\*
- 11: gb\_sy.\*
- 12: gb\_un.\*
- 13: gb\_vi.\*
- 14: gb\_htg.\*
- 15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1669	100.0	1669	6	AR651454
2	1669	100.0	1669	8	AF221636
3	1659	99.4	2668	6	AX405771
4	1659	99.4	2842	6	AX405770
5	1655	99.2	2778	8	AK000290
6	1648.4	98.8	4523	6	AR631281
7	1648.4	98.8	4523	6	AR631281
8	1505	90.2	3125	8	BC030688
9	1501	89.9	4676	6	AR631287
10	1501	89.9	4676	6	AX608743
11	1487.4	89.1	2161	6	BD157001
12	1487.4	89.1	2161	6	AX878058
13	1487.4	89.1	2161	8	AK027826
14	1354	81.1	3127	8	AF221634
15	1353.4	81.1	3143	6	AX354793
16	1349.6	80.9	4829	6	AR631283
17	1349.6	80.9	4829	6	AX608735
18	1347	80.7	3120	6	AR651452

19	1347	80.7	4535	8	BC040203	BC040203 Homo sapi
20	1344	80.5	3106	6	AX342633	AX342633 Sequence
21	1343.6	80.5	4685	6	AR631288	AR631288 Sequence
22	1343.6	80.5	4685	6	AX608745	AX608745 Sequence
23	1337	80.1	2797	6	AR448400	AR448400 Sequence
24	1250.6	74.9	2510	6	AX338497	AX338497 Sequence
25	1234	73.9	3030	8	AY354202	AY354202 Homo sapi
26	1087.4	65.2	2649	6	AX354795	AX354795 Sequence
27	1085.8	65.1	2656	8	AY172659	AY172659 Homo sapi
28	1085.8	65.1	2671	6	AR631278	AR631278 Sequence
29	1085.8	65.1	2671	6	AX608725	AX608725 Sequence
30	924.6	55.4	4799	9	BC043124	BC043124 Mus muscu
31	924.6	55.4	4799	9	BC059222	BC059222 Mus muscu
32	881	52.8	1197	8	AR651453	AR651453 Sequence
33	881	52.8	1197	8	AF221635	AF221635 Homo sapi
34	765.2	45.8	4309	6	AR631284	AR631284 Sequence
35	765.2	45.8	4309	6	AX608737	AX608737 Sequence
36	636.4	38.1	873	6	AX083130	AX083130 Sequence
37	605.6	36.3	183341	14	AC161443	AC161443 Mus muscu
C. 38	589.2	35.3	194188	9	AC163387	AC163387 Mus muscu
39	578	34.6	753	6	AX524935	AX524935 Sequence
40	524	31.4	735	6	AX524942	AX524942 Sequence
C 41	523.6	31.4	587	6	BD154822	BD154822 Primer fo
C 42	523.6	31.4	587	6	AX874760	AX874760 Sequence
43	493.4	29.6	823	6	AR263993	AR263993 Sequence
44	446.2	26.7	1002	6	AX524939	AX524939 Sequence
45	440	26.4	2098	6	CQ721272	CQ721272 Sequence

ALIGNMENTS

RESULT 1  
AR651454  
LOCUS AR651454 1669 bp DNA linear PAT 20-APR-2005  
DEFINITION Sequence 6 from patent US 6881564.  
ACCESSION AR651454  
VERSION AR651454.1 GI:62795940  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1669)  
AUTHORS Abbott, C.A. and Gorrell, M.D.  
TITLE Dipeptidyl peptidases  
JOURNAL Patent: US 6881564-A 6 19-APR-2005;  
The University of Sydney; Sydney;  
AUX;  
FEATURES  
source Location/Qualifiers  
1..1669  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN	Query Match	Best Local Similarity	Score	DB 6;	Length	1669;
	Matches 1669;	Conservative	0;	Mismatches	0;	Indels
						Gaps
						0;
Qy	1	AACAGGTACAGCAATCCTTAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGA	60			
Db	1	AACAGGTACAGCAATCCTTAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGA	60			
Qy	61	AGGAAGGATCATAGATGTCATAGATAAGCACTTAATTCACCTTTTGAGATTCCTATTTGA	120			
Db	61	AGGAAGGATCATAGATGTCATAGATAAGCACTTAATTCACCTTTTGAGATTCCTATTTGA	120			
Qy	121	AGGAGTTGAATATATTCGCCAGAGCTGGATCGGAGGAAATATGCTTGGTCCAT	180			
Db	121	AGGAGTTGAATATATTCGCCAGAGCTGGATCGGAGGAAATATGCTTGGTCCAT	180			
Qy	181	CCTACTAGTCGCTCCAGACTCGCTACAGATAGTGTGATCTACCTGGAATTTAT	240			
Db	181	CCTACTAGTCGCTCCAGACTCGCTACAGATAGTGTGATCTACCTGGAATTTAT	240			

Qy	241	CCAGTAGAAGATGATGTTATTCGAAAGGAGAGACTCATTTGAGTCACTGCTGATCTGT	300
Db	241	CCAGTAGAAGATGATGTTATTCGAAAGGAGAGACTCATTTGAGTCACTGCTGATCTGT	300
Qy	301	GAGCCACTAATATCTATGAAAGAAACAACAGACATCTGGATAAATATCCATGACATCT	360
Db	301	GAGCCACTAATATCTATGAAAGAAACAACAGACATCTGGATAAATATCCATGACATCT	360
Qy	361	TGATGTTTTTCCCAAGTTCAGAAAGAGAAATGAGTTATTTTTTGCCTCTGAAATGCAA	420
Db	361	TGATGTTTTTCCCAAGTTCAGAAAGAGAAATGAGTTATTTTTTGCCTCTGAAATGCAA	420
Qy	421	AACAGGTTTCCGTCATTTATACAAAATACATCTATTTTAAAGGAAAGCAAAATATAACG	480
Db	421	AACAGGTTTCCGTCATTTATACAAAATACATCTATTTTAAAGGAAAGCAAAATATAACG	480
Qy	481	ATCCAGTGTGGGCTGCCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC	540
Db	481	ATCCAGTGTGGGCTGCCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC	540
Qy	541	AATTACAGTGTGATGGAAGTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGA	600
Db	541	AATTACAGTGTGATGGAAGTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGA	600
Qy	601	AGTCAGAGGCTGATATTTTAAAGGACCAAGACCTCCCTTTAGAGCATCACTGTA	660
Db	601	AGTCAGAGGCTGATATTTTAAAGGACCAAGACCTCCCTTTAGAGCATCACTGTA	660
Qy	661	CGTAGTCAGTTACGTAATCTCGAGAGAGTGACAAAGCTGACGACCGTGTACTACAA	720
Db	661	CGTAGTCAGTTACGTAATCTCGAGAGAGTGACAAAGCTGACGACCGTGTACTACAA	720
Qy	721	TTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAGAA	780
Db	721	TTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAGAA	780
Qy	781	TCCACACTGTGTCTCCCTTTTAAAGCTATCAAGTCTCAAGTACGACCACTTGCAAAAC	840
Db	781	TCCACACTGTGTCTCCCTTTTAAAGCTATCAAGTCTCAAGTACGACCACTTGCAAAAC	840
Qy	841	AAAGGAATTTGGGCGCCACATTTGGATTCAGCAGTCTCTCTGACTATATCTCTCC	900
Db	841	AAAGGAATTTGGGCGCCACATTTGGATTCAGCAGTCTCTCTGACTATATCTCTCC	900
Qy	901	AGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCC	960
Db	901	AGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCC	960
Qy	961	TCATGATCTACAGCTGGAAGATATCTCTGCTGCTGTTTATATATGTTGCTCTCA	1020
Db	961	TCATGATCTACAGCTGGAAGATATCTCTGCTGCTGTTTATATATGTTGCTCTCA	1020
Qy	1021	GGTGTATGCTGGGCGCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGA	1080
Db	1021	GGTGTATGCTGGGCGCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGA	1080
Qy	1081	ACGTTATGAGGTCACCTGACAGAAAGAGGCTATTTACTTAGGATCTGTGGCAT	1140
Db	1081	ACGTTATGAGGTCACCTGACAGAAAGAGGCTATTTACTTAGGATCTGTGGCAT	1140
Qy	1141	GCAAGCAGAAAGTTCCCTCTGAAACCAAAATGTTTACTGCTCTTACATGGTTCTCGGA	1200
Db	1141	GCAAGCAGAAAGTTCCCTCTGAAACCAAAATGTTTACTGCTCTTACATGGTTCTCGGA	1200
Qy	1201	TGAAATGTCATTTTGGCAATACAGATATTTACTGAGTTTCTTTAGTGGGCTGGAA	1260
Db	1201	TGAAATGTCATTTTGGCAATACAGATATTTACTGAGTTTCTTTAGTGGGCTGGAA	1260
Qy	1261	GCAATATGATTTTACAGATCTATCTCTCAGAGAGACACAGATATGAGTTCTGAATCGGG	1320
Db	1261	GCAATATGATTTTACAGATCTATCTCTCAGAGAGACACAGATATGAGTTCTGAATCGGG	1320
Qy	1321	AGAACATTATGAATCCTCATCTTTTGCACCTTCAAGAAACCTTGATCAGCTATTC	1380
Db	1321	AGAACATTATGAATCCTCATCTTTTGCACCTTCAAGAAACCTTGATCAGCTATTC	1380
Qy	1381	TGCTCTAAAAGTATATAATTTTGACCTGTGTAGAACTCTCTGTGATACACTGGCTATTT	1440
Db	1381	TGCTCTAAAAGTATATAATTTTGACCTGTGTAGAACTCTCTGTGATACACTGGCTATTT	1440
Qy	1441	AACAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCATCTTTGATACC	1500
Db	1441	AACAAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCATCTTTGATACC	1500
Qy	1501	TGCCATGTAACTACTCTCTGAAATAAATGTGTGTCATGCGGGGTCTACGGTTGT	1560
Db	1501	TGCCATGTAACTACTCTCTGAAATAAATGTGTGTCATGCGGGGTCTACGGTTGT	1560
Qy	1561	GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATCATATTCCTGAGA	1620
Db	1561	GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATCATATTCCTGAGA	1620
Qy	1621	GACCCAGCAATACCAAGAATTTACTTAAAAAATAAAAAAATAAAAAA 1669	
Db	1621	GACCCAGCAATACCAAGAATTTACTTAAAAAATAAAAAAATAAAAAA 1669	
RESULT 2			
AF221636			
LOCUS Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, partial cds,			
DEFINITION Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, partial cds,			
ACCESSION AF221636			
VERSION AF221636.1 GI:11095191			
KEYWORDS			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominiidae; Homo.			
REFERENCE 1 (bases 1 to 1669)			
AUTHORS Abbott,C.A., Yu,D.M., Woolliatt,E., Sutherland,G.R., McCaughan,G.W.			
and Gorrell,M.D.			
TITLE Cloning, expression and chromosomal localization of a novel human			
dipeptidyl peptidase (DPP) IV homolog, DPP8			
Eur. J. Biochem. 267 (20), 6140-6150 (2000)			
REFERENCE 2 (bases 1 to 1669)			
PUBMED 11012666			
AUTHORS Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.			
TITLE Direct Submission			
JOURNAL Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver			
Centre, Centenary Institute of Cell Biology and Cancer Medicine,			
Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia			
FEATURES			
SOURCE			
1..1669			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon.9606"			
/chromosome="15"			
/map="15q22"			
/clone="18"			
/tissue_type="placenta"			
/clone_lib="lambda triplexExTM: Clontech"			
<1..1669			
/gene="DPP8"			
<1..1399			
/gene="DPP8"			
/note="serine protease; member of prolyl oligopeptidase S9b			
family belonging to enzyme clan SC; isoform missing the			
catalytic serine required for enzyme activity; additional			
splice variants of this mRNA are found in GenBank			
Accession Numbers AF221634, AF221635 and AF221637"			
/codon_start=2			
/product="dipeptidyl peptidase 8"			
/protein_id="AAG29768.1"			



		/db_xref="GI:11095192"	
		/translation="TGTANKVTFKMSIMIDABGRILIDVIDKELIQFBIIFEGVEY IARAGTPEKAYMSILLDSQRIQLVILISPELPIVEDDWMERQLTIESVPSVTP LIIVEETDIIWINIHDPHFVFPQSHEEIEIFASECKTGFRLKYKLTSLIKESKYKR SSGLPAPSDKCPKEIEIATISGEWEVLGRHSNIOVDEVRRLVYFEGTKDSPLHH LYVVSVPGEVTRLDRCYSHSCCISOCHDFLSKYSNOKNPHCVSLYKLSPEDDP TKTKFENATILDSAGLDPDYDYPPEIFPFESTGFTLYGMLYKPHLDQPKKYPVLV IYGPQVATAGASVTLWIFYDTGYTERYHGHDPQNEOGYILGVSAMQAEKFPSEPNRL LLHGLDENVHPAHTSILLSLFLVRAGKPYDLQYPOERHSIRVPSSGEHYELHLHY LOENLASRIAAALKVI"	
ORIGIN			
Query Match		100.0%; Score 1669; DB 8; Length 1669;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATCTGA	60
DB	1	AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATCTGA	60
QY	61	AGGAAGGATCATAGATGTCATAGATAAGGAATAATTCACCTTTTGAGATTCCTATTGA	120
DB	61	AGGAAGGATCATAGATGTCATAGATAAGGAATAATTCACCTTTTGAGATTCCTATTGA	120
QY	121	AGGATTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATAATGCTTGGTCCAT	180
DB	121	AGGATTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATAATGCTTGGTCCAT	180
QY	181	CCTACTAGATCGCTCCGAGCTCGCTACAGATAGTGTGATCTCACCTGAATATTATTAT	240
DB	181	CCTACTAGATCGCTCCGAGCTCGCTACAGATAGTGTGATCTCACCTGAATATTATTAT	240
QY	241	CCGAGTAAAGATGATGTTATGGAAGGACAGAGACTCAATGAGTCAGTGCCTGATTCCTGT	300
DB	241	CCGAGTAAAGATGATGTTATGGAAGGACAGAGACTCAATGAGTCAGTGCCTGATTCCTGT	300
QY	301	GAGCCCACTAATTTATCTATGAAGAAACAAACAGATCTGGATAAATATCCATGACATCTT	360
DB	301	GAGCCCACTAATTTATCTATGAAGAAACAAACAGATCTGGATAAATATCCATGACATCTT	360
QY	361	TCATGTTTTTCCCAAGTCACGAGAGGAATGAGTTATTTTGGCTCTGGAATGCAA	420
DB	361	TCATGTTTTTCCCAAGTCACGAGAGGAATGAGTTATTTTGGCTCTGGAATGCAA	420
QY	421	AACAGGTTTCGGTCAATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAAAG	480
DB	421	AACAGGTTTCGGTCAATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAAAG	480
QY	481	ATCCAGTGTGGGCTGCTCTCCAAAGTATTTCAAGTGTCTATCAAGAGGAGATAGC	540
DB	481	ATCCAGTGTGGGCTGCTCTCCAAAGTATTTCAAGTGTCTATCAAGAGGAGATAGC	540
QY	541	AATTACAGTGTGTAATGGAGATTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA	600
DB	541	AATTACAGTGTGTAATGGAGATTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA	600
QY	601	AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTA	660
DB	601	AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTA	660
QY	661	CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGTACTACCA	720
DB	661	CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGTACTACCA	720
QY	721	TTCTTGTGTCATCAGTCAGACTGTGACTTTTATAAGTAAAGTATAGTAAACCAAGAGAA	780
DB	721	TTCTTGTGTCATCAGTCAGACTGTGACTTTTATAAGTAAAGTATAGTAAACCAAGAGAA	780
QY	781	TCCACACTGTGCTCCCTTTTACAGCTATCAAGTCTCAAGTACCCCACTTCGCAAAAC	840
DB	781	TCCACACTGTGCTCCCTTTTACAGCTATCAAGTCTCAAGTACCCCACTTCGCAAAAC	840
QY	841	AAAGGAATTTTGGGCCACCACTTTTGGATTCAGCAGGCTCTCTCTGACTATATCTCCTCC	900
DB	841	AAAGGAATTTTGGGCCACCACTTTTGGATTCAGCAGGCTCTCTCTGACTATATCTCCTCC	900
QY	901	AGAAATTTTCTTTTGGAAAGTACTACTGATTACATTTGATGGATGCTCTACAAAGCC	960
DB	901	AGAAATTTTCTTTTGGAAAGTACTACTGATTACATTTGATGGATGCTCTACAAAGCC	960
QY	961	TCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCA	1020
DB	961	TCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCA	1020
QY	1021	GGTGTCTATTGCTGGGGCCCACTCTCTGTGGATCTCTATGATACAGGATACACGA	1080
DB	1021	GGTGTCTATTGCTGGGGCCCACTCTCTGTGGATCTCTATGATACAGGATACACGA	1080
QY	1081	ACGTTATATGGGTCACTCCCTGACCCAGAAATGAAACAGGCTATTACTTAGGATCTGTGCCAT	1140
DB	1081	ACGTTATATGGGTCACTCCCTGACCCAGAAATGAAACAGGCTATTACTTAGGATCTGTGCCAT	1140
QY	1141	GCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGGA	1200
DB	1141	GCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGGA	1200
QY	1201	TGAGAAATGTCATTTTGCACATACAGATATATCTAGAGTTTCTTAGAGGCTGGAAA	1260
DB	1201	TGAGAAATGTCATTTTGCACATACAGATATATCTAGAGTTTCTTAGAGGCTGGAAA	1260
QY	1261	GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGAGTTCTTGAATCGGG	1320
DB	1261	GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGAGTTCTTGAATCGGG	1320
QY	1321	AGAACATTTAGAACTGTCATCTTTTGCACACTACCTTCAAGAAAAACCTTGGATCACTATTGC	1380
DB	1321	AGAACATTTAGAACTGTCATCTTTTGCACACTACCTTCAAGAAAAACCTTGGATCACTATTGC	1380
QY	1381	TGCTCTAAAAGTGAATATAATTTTGAACCTGTGTAGAACTCTCTGTGTATATACATCGGCTATTT	1440
DB	1381	TGCTCTAAAAGTGAATATAATTTTGAACCTGTGTAGAACTCTCTGTGTATATACATCGGCTATTT	1440
QY	1441	AACCAAAATGAGGAGTTTAAATCAACAGAAAAACAGAAATGATCATCATTTTGCATACC	1500
DB	1441	AACCAAAATGAGGAGTTTAAATCAACAGAAAAACAGAAATGATCATCATTTTGCATACC	1500
QY	1501	TGCCATGTAAACATCTACTCTCTGAAAAATAATGTGTGTCCTACGAGGGGTCTACGGTTTGT	1560
DB	1501	TGCCATGTAAACATCTACTCTCTGAAAAATAATGTGTGTCCTACGAGGGGTCTACGGTTTGT	1560
QY	1561	GGTAGTAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTCCTGAGA	1620
DB	1561	GGTAGTAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTCCTGAGA	1620
QY	1621	GACCCAGCAATACCAATGAAATTTACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTT	1669
DB	1621	GACCCAGCAATACCAATGAAATTTACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTT	1669
RESULT 3		;	
AX405771		Sequence 186 from Patent WO0222660.	
LOCUS		2668 bp DNA linear PAT 14-JUN-2002	
DEFINITION		Sequence 186 from Patent WO0222660.	
ACCESSION		AX405771	
VERSION		AX405771.1 GI:21438981	
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Homo sapiens	
REFERENCE		Tang, X.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F., Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T. Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Tang, X.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F., Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.	
TITLE		Novel nucleic acids and polypeptides	
JOURNAL		Patent: WO 0222660-A 186 21-MAR-2002;	

FEATURES	HYSEQ, INC. (US)	
	Location/Qualifiers	
source	1..2668	
	/organism="Homo sapiens"	
CDS	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
	234..2408	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD34878.1"	
	/db_xref="GI:21439982"	
	/translation="WAAAMETQLGVIFETADCEENIESQDRPKLEPFVVERYSWQ LKLLADRKHYGVNMAKAPHDPMFKVKNPDGPHSDRIYVLLAMSENRENTLPFSEI PTXIRAAVLMSKPLDLFQQQLRNLAVETSCPNIRMDPKLCPADDMWLFTHSN DTWISNIVREERRLTYYHNELANMEDSASGAVFVLOEEDRYSYVMCPKFAET PGGKILRLYEENDESEVELIHVTSPLMETRADSPRYPKGTGNPKPTFKMSIMI DAEGLIDVLDKELQPELIFEGVEVIAKAGHTPBGKYAWSLIDRSOTRLQIVLIS PLFVPEVDUMERORLIESVPDSVTBLIYESTDMINIHDIHVPFROSHERIEF IFASCKTGFRHLYXITSLIKESRYKSSGGLPAPSDPKPIKEETATPSGHEVLGR HSGNQDEVRRLVTFEGTKOSPLEHLLYVSVYNFQEVRLDRGYSKSCISQHCD FTISKSNQKPHCVSLYKLSPEDDPTCKIKFPMATILDSAGLPEDYTPPEIFSFES TTGFLYGMLYKPHDLQPKKYPVLFYIGGPQVIAAGAEVILWIFIDYGTIERHGH PDNSGGYTLGSVAQAEKFPSEPNRLLLHGLFDENVHFAHTSILLSLFLVRAGRPYD LQIYPOERHSIRVPESGHEHLLHLYQENLGSRIAALKVI"	
ORIGIN		
Query Match		99.4%; Score 1659; DB 6; Length 2668;
Best Local Similarity		100.0%; Pred. No. 0;
Matches 1659; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAAGATGTGAGAAATGATGATGATGCTGA	60
Db	1010 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAAGATGTGAGAAATGATGATGATGCTGA	1069
Qy	61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTTCAGAGATTCATTTGA	120
Db	1070 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTTCAGAGATTCATTTGA	1129
Qy	121 AGGAGTTGAATATTTGGCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCAT	180
Db	1130 AGGAGTTGAATATTTGGCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCAT	1189
Qy	181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCCTGAATTTAT	240
Db	1190 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCCTGAATTTAT	1249
Qy	241 CCCAGTAGAGATGATGTTTGGAAAGCGAGACTCAATTGAGTCAGTCTGATCTGT	300
Db	1250 CCCAGTAGAGATGATGTTTGGAAAGCGAGACTCAATTGAGTCAGTCTGATCTGT	1309
Qy	301 GAGCCACTAATTTATCTATGAAGAACACAGACATCTGGAATTAATCCATCAGATCTT	360
Db	1310 GAGCCACTAATTTATCTATGAAGAACACAGACATCTGGAATTAATCCATCAGATCTT	1369
Qy	361 TCATGTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTCCTCTGAAATGCAA	420
Db	1370 TCATGTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTCCTCTGAAATGCAA	1429
Qy	421 AACAGGTTCCGTCATTTTACAAAATTCATCTATTTTAAAGGAAAGCAAAATATAACG	480
Db	1430 AACAGGTTCCGTCATTTTACAAAATTCATCTATTTTAAAGGAAAGCAAAATATAACG	1489
Qy	481 ATCCAGTGTGGGCTCGCTGCTCCAAAGTGAATTCAGGTGTCCTATCAAGAGGAGATAGC	540
Db	1490 ATCCAGTGTGGGCTCGCTGCTCCAAAGTGAATTCAGGTGTCCTATCAAGAGGAGATAGC	1549
Qy	541 AATTACAGTGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	600
Db	1550 AATTACAGTGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	1609
Qy	601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA	660
Db	1610 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA	1669
AX405770		2842 bp DNA linear PAT 14-JUN-2002

DEFINITION	Sequence 185 from Patent WO0222660.	
ACCESSION	AX405770	
VERSION	AX405770.1	GI:21438979
KEYWORDS	Hom sapiens (human)	
SOURCE	Hom sapiens (human)	
ORGANISM	Hom sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F., Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.	
TITLE	Novel nucleic acids and polypeptides	
JOURNAL	Patent: WO 0222660-A 185 21-MAR-2002;	
HYSEQ, INC. (US)	HYSEQ, INC. (US)	
FEATURES	Location/Qualifiers	
source	1..2842	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
CDS	234..2582	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD34877.1"	
	/db_xref="GI:21438980"	
	/translations="MAAMETBQLGVEIFETADCEENIESQDRPKLBPFFYVYSMSQ LKLLADTRRYHGMKAPHDPMFKRNDPDGHSRDIYYLWAMSGNRENTLFYSI PKNINRAALVHWKPLLDFOATLDYGMYSREELLRRKRI GTVGIASVDYHQSG TFLPQAGSGLVHKDGGPGQFTQPLRNPLVETSPNIRMDPKLCPADPDWIAFHSN DIWLSNIVTEERRLTYVHNLNNEEDARSAGVATFVLQEFPRYSGYWCPKAETI PSQGIILRIYBENDESEVEI IHVTSFMLETRRADSFYPKTGANPKVTFKXSEIMI DAEGRIIDVDEKLIQFPEILFEGVEYIARAGMTPEGYAWSILLDRSQRLOVLIS PELFIPVEDDMERQRLIESVDSVTPLI IYEETD IWINIHDFHVPQSHBEEIEF IPASECTGVRHLVYKITSILKESKYKRSGGLPAPSDPKCPKEEIAITSGEWELGR HGSNIQVDEVRRLVYREGTKDSPLEHLVYVSVNAPGSEVTRLDYRGYSHSCISOHCD FFTGISQNKPNHCVSILYKLSPEDDPTCKTFWATILDSAGELPDPPEIPFESFH TTGTLTKYLDKPHDLOPKKYPTVLFYIGGPQVAIAGAPVTLWIFDYDTGTERYMGH PDNEQGYILGSVAMQAEKFPSEPNRLHLLHGLFDENVHFAHTSILLFSLVRAGKPYD LQIYQERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI"	
ORIGIN		
Query Match	99.4%; Score 1659; DB 6; Length 2842;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1659; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGCTGA	60
DB	1184 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGCTGA	1243
QY	61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAAATTCACACCTTTTGAGATTCTATTGA	120
DB	1244 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAAATTCACACCTTTTGAGATTCTATTGA	1303
QY	121 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAAATTCACACCTTTTGAGATTCTATTGA	180
DB	1304 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAAATTCACACCTTTTGAGATTCTATTGA	1363
QY	181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATATTAT	240
DB	1364 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATATTAT	1423
QY	241 CCCAGTGAAGATGATGTTATGGAAGGAGAGACTCATTTAGTGCAGTGCCTGATTCTGT	300
DB	1424 CCCAGTGAAGATGATGTTATGGAAGGAGAGACTCATTTAGTGCAGTGCCTGATTCTGT	1483
QY	301 GACGCCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT	360
DB	1484 GACGCCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT	1543
QY	361 TCATGTTTTTCCCCAAGTCAAGAGGAAATGAGTTATTTTGGCTCTGGAATGCAA	420
DB	1544 TCATGTTTTTCCCCAAGTCAAGAGGAAATGAGTTATTTTGGCTCTGGAATGCAA	1603

|||||  
2684 TGCCATGTAACATCTACTCTCTGAAATAAATGTGTGCGCATGCGGGGTCTACGGTTGT 2743  
|||||  
1561 GGTAGTAATCTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 1620  
|||||  
2744 GGTAGTAATCTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 2803  
|||||  
1621 GACCCAGCAATACCAATGAATTAATCTAAAAAATAAAAA 1659  
|||||  
2804 GACCCAGCAATACCAATGAATTAATCTAAAAAATAAAAA 2842  
|||||  
RESULT 5  
AK000290 2778 bp mRNA linear PRI 13-SEP-2003  
LOCUS Homo sapiens cDNA FLJ20283 f16, clone HEP04088.  
DEFINITION  
ACCESSION AK000290  
VERSION AK000290.1 GI:7020272  
KEYWORDS oligo capping; f16 (full insert sequence).  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1  
Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,  
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T.,  
Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,  
Nakamura,Y., Isogai,T. and Sugano,S.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2778)  
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,  
Shibahara,T., Tanaka,T. and Nakamura,Y.  
Direct Submission  
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,  
University of Tokyo, Department of Virology; Shirokane-dai, 4-6-1,  
Minato-ku, Tokyo 108-8639, Japan (E-mail:flcdna@ims.u-tokyo.ac.jp,  
Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing; Research Association for Biotechnology; cDNA library  
construction, 5'- & 3'-end one pass sequencing; Department of  
Virology and Human Genome Center, Institute of Medical Science,  
University of Tokyo (partly supported by Science and Technology  
Agency). Location/Qualifiers  
FEATURES  
source 1.  
2778  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_line="HEP04088"  
/cell\_type="HepG2"  
/clone\_lib="HEP"  
/notes="cloning vector pME18SFL3"  
163..1572  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAA91059.1"  
/db\_xref="GI:7020273"  
/translation="MAAAMETQLGVIFETADCEENIESODRPLEFFVVERYSWSQ  
LKLLADTRKYGYMMAKAPHDPMFKRNDPGPHSDRIYYLWMSGENRNTILFYSEI  
PKTNRAAIVLMSWKPLLDLFOATLDYGMYSREELRERKRIGTGIASDYHQSSG  
TFLFQAGSGIYHKVGGPQGTQQPLRENVLVETCPNIRMDPKLCPADPOMIAPHSN  
DIWISNIVREERLTYVHNELANMEEDARGAVATFVLOEEDFDRYSGYWPCPKAETI  
PSGGKILRLIYEENDESEVEI IHVTSPLMETRADSPRYKPTGPANPKVTFPKAETI  
DASGRIDVIDKELIQFPEILFEGVEYIARAGWTPEGKIAMSILLDRSQRLQVLIS  
PELFIPIVDVMEORLIESVPDSVTPLIYEETTDWINIHDIHFVFPQSHBEEIEF  
IFASECKTGRPLKYKITSI"  
ORIGIN  
Query Match 99.2%; Score 1655; DB 8; Length 2778;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Indels 1; Gaps 1;  
Matches 1666; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 1 AACAGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATATATGATGCTGA 60  
Db 1113 AACAGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATATATGATGCTGA 1172  
Qy 61 AGGAAGGATCATAGATGTCATAGATAAGGAACTAAATCAACCTTTTGAAGATTCATTGGA 120  
Db 1173 AGGAAGGATCATAGATGTCATAGATAAGGAACTAAATCAACCTTTTGAAGATTCATTGGA 1232  
Qy 121 AGGAGTTGAATATATTGCGAGAGCTGGATGAGCTCCTGAGGGGAAAATATGCTTGGTCCAT 180  
Db 1233 AGGAGTTGAATATATTGCGAGAGCTGGATGAGCTCCTGAGGGGAAAATATGCTTGGTCCAT 1292  
Qy 181 CCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTGTGATCTCAGCTCACTCAATATTATAT 240  
Db 1293 CCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTGTGATCTCAGCTCACTCAATATTATAT 1352  
Qy 241 CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATTCAGTCAGTGCCTGATTCCTGT 300  
Db 1353 CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATTCAGTCAGTGCCTGATTCCTGT 1412  
Qy 301 GAGCCACTAATTTCTATGAAAGCAACAGACATCTGGATAAATATCCATGACATCTT 360  
Db 1413 GAGCCACTAATTTCTATGAAAGCAACAGACATCTGGATAAATATCCATGACATCTT 1472  
Qy 361 TCATGTTTTTCCCAAGTCAGAGAGAAATTTGAGTTATTTTTCCTCTGAAATGCA 420  
Db 1473 TCATGTTTTTCCCAAGTCAGAGAGAAATTTGAGTTATTTTTCCTCTGAAATGCA 1532  
Qy 421 AACAGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAAAATATAACG 480  
Db 1533 AACAGTTTCCGTCATTTATACAAAATTTACATCTA-TTTAAAGGAAGCAAAATATAACG 1591  
Qy 481 ATCCAGTGGTGGCTCCCTGCTCCCAAGTATTTCAAGTGTCTTCAAGAGGAGATAGC 540  
Db 1592 ATCCAGTGGTGGCTCCCTGCTCCCAAGTATTTCAAGTGTCTTCAAGAGGAGATAGC 1651  
Qy 541 AATTACAGTGGTGAATGGAAAGTTCTTGGCCGCGATGGATCTAATATCCAAAGTTGATGA 600  
Db 1652 AATTACAGTGGTGAATGGAAAGTTCTTGGCCGCGATGGATCTAATATCCAAAGTTGATGA 1711  
Qy 601 AGTCAGAGGCTCGTATATTTTGAAGGCAACAAAGACTCCCTTTTGAAGCATCACCTGTA 660  
Db 1712 AGTCAGAGGCTCGTATATTTTGAAGGCAACAAAGACTCCCTTTTGAAGCATCACCTGTA 1771  
Qy 661 CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 720  
Db 1772 CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 1831  
Qy 721 TTCTTGTGTCATCAGTCAGCACTGTGACTTTCTTTATAAGTAAATAGTATAGTAAACCAAGAA 780  
Db 1832 TTCTTGTGTCATCAGTCAGCACTGTGACTTTCTTTATAAGTAAATAGTATAGTAAACCAAGAA 1891  
Qy 781 TCCACACTGTGTCTCCCTTTACAAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGCAAAAC 840  
Db 1892 TCCACACTGTGTCTCCCTTTACAAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGCAAAAC 1951  
Qy 841 AAAGGAATTTTGGCCCAACCAATTTGGATTTCAGCAGTCTCTCTCTGCTACTACTCTCTCC 900  
Db 1952 AAAGGAATTTTGGCCCAACCAATTTGGATTTCAGCAGTCTCTCTCTGCTACTACTCTCTCC 2011  
Qy 901 AGAATTTTCTCTTTTGAAGTACTCTGATTTCAATTTGATGATGATGATGATGATGATGATGAT 960  
Db 2012 AGAATTTTCTCTTTTGAAGTACTCTGATTTCAATTTGATGATGATGATGATGATGATGATGAT 2071  
Qy 961 TCATGATCTACAGCTGGAGAAAGAAATATCTCTACTGTGTGTCATATATGATGATGATGATGATGAT 1020  
Db 2072 TCATGATCTACAGCTGGAGAAAGAAATATCTCTACTGTGTGTCATATATGATGATGATGATGATGAT 2131  
Qy 1021 GGTGTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTCTCTATGATACAGATACACGGA 1080

Db	2132	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	2191
Qy	1081	ACGTTATATGGGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT	1140
Db	2192	ACGTTATATGGGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT	2251
Qy	1141	GCAGCAGAAAATGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGA	1200
Db	2252	GCAGCAGAAAATGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGA	2311
Qy	1201	TGAGAAATGTTCCATTTTGCACATACCAAGTATATTAATCTAGTGTATTTTACTGAGGCTCGAAA	1260
Db	2312	TGAGAAATGTTCCATTTTGCACATACCAAGTATATTAATCTAGTGTATTTTACTGAGGCTCGAAA	2371
Qy	1261	GCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCCTGAATCGGG	1320
Db	2372	GCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCCTGAATCGGG	2431
Qy	1321	AGAACATTAATGATCTCTTTTGCACATACCTTCAAGAAAACTTGGGATCAAGTATTC	1380
Db	2432	AGAACATTAATGATCTCTTTTGCACATACCTTCAAGAAAACTTGGGATCAAGTATTC	2491
Qy	1381	TGCTCTAAAGTGATATTAATTTGACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTT	1440
Db	2492	TGCTCTAAAGTGATATTAATTTGACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTT	2551
Qy	1441	AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCACTTTTGATACC	1500
Db	2552	AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCACTTTTGATACC	2611
Qy	1501	TGCAATGATTAACATCTACTCTCTGAAATTAATGTTGGTGCATGCAAGGGTCTACGGTTGT	1560
Db	2612	TGCAATGATTAACATCTACTCTCTGAAATTAATGTTGGTGCATGCAAGGGTCTACGGTTGT	2671
Qy	1561	GGTAGTAATCTAATACCTTTAACCACCATGCTCAAAATCAAAATGATACATATTCCTGAGA	1620
Db	2672	GGTAGTAATCTAATACCTTTAACCACCATGCTCAAAATCAAAATGATACATATTCCTGAGA	2731
Qy	1621	GACCCAGCAATACCAATAGAAATTTACTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA	1667
Db	2732	GACCCAGCAATACCAATAGAAATTTACTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA	2778
RESULT 6			
LOCUS	AR631281	4523 bp	DNA linear PAT 14-FEB-2005
DEFINITION	Sequence 8 from patent US 6844180.		
ACCESSION	AR631281		
VERSION	AR631281.1	GI:59770924	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 4523)		
AUTHORS	Oj, S., Akinsanya K.O., Riviere, P.J.M. and Junien, J.-L.		
TITLE	Serine protease genes related to DPPIV		
JOURNAL	Patent: US 6844180-A 8 18-JAN-2005;		
	Ferring BV;		
FEATURES	NLX;		
source	Location/Qualifiers		
	1..4523		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
	Query Match 98.8%; Score 1648.4; DB 6; Length 4523;		
	Best Local Similarity 99.7%; Pred. No. 0;		
	Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;		
Qy	1	AACAGGTACAGCAATCTTAAGTCACTTTTAAAGATGTACAGAAATATGATGATGCTGA	60
Db	1164	AACAGGTACAGCAATCTTAAGTCACTTTTAAAGATGTACAGAAATATGATGATGCTGA	1223

Qy	61	AGGAGGATCATAGATGTCATAGATAAGGAACTAATTCACACTTTTGAGATTCTATTGGA	120
Db	1224	AGGAGGATCATAGATGTCATAGATAAGGAACTAATTCACACTTTTGAGATTCTATTGGA	1283
Qy	121	AGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCCTCGAGGAAAAATATGCTTGGTCCAT	180
Db	1284	AGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCCTCGAGGAAAAATATGCTTGGTCCAT	1343
Qy	181	CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCTCACTGAATATTATAT	240
Db	1344	CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCTCACTGAATATTATAT	1403
Qy	241	CCCAGTAGAAGATGATGTTATGGAAGGCGAGAGCTCATTTGAGTCAGTCAGTCGCTGATCTGT	300
Db	1404	CCCAGTAGAAGATGATGTTATGGAAGGCGAGAGCTCATTTGAGTCAGTCAGTCGCTGATCTGT	1463
Qy	301	GAGCCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT	360
Db	1464	GAGCCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT	1523
Qy	361	TCATGTTTTTCCCAAGTCA CGAAGAGGAAAATTGAGTTTATTTTGGCTCTGGAATGCAA	420
Db	1524	TCATGTTTTTCCCAAGTCA CGAAGAGGAAAATTGAGTTTATTTTGGCTCTGGAATGCAA	1583
Qy	421	AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAACG	480
Db	1584	AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAACG	1643
Qy	481	ATCCAGTGTGGGCTGCCCTGCCAAGTGATTTCAAGTGTCCTATCAAAAGAGGAGATAGC	540
Db	1644	ATCCAGTGTGGGCTGCCCTGCCAAGTGATTTCAAGTGTCCTATCAAAAGAGGAGATAGC	1703
Qy	541	AAATTCACAGTGTGAAATGGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	600
Db	1704	AAATTCACAGTGTGAAATGGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	1763
Qy	601	AGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGACCTCCCTTTAGAGCATCACTGTGA	660
Db	1764	AGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGACCTCCCTTTAGAGCATCACTGTGA	1823
Qy	661	CGTAGTCAGTTACGTAATAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	720
Db	1824	CGTAGTCAGTTACGTAATAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA	1883
Qy	721	TTCTTGTCGATCAGTCAGCACTGTGACTTTCTTTATAAGTAAGTATAGTATACCAAGAGAA	780
Db	1884	TTCTTGTCGATCAGTCAGCACTGTGACTTTCTTTATAAGTAAGTATAGTATACCAAGAGAA	1943
Qy	781	TCCACACTGTGTGTCCTCTTTACAGCTATCAAGTCTCTGAGATGACCCAACTTGCAGAAC	840
Db	1944	TCCACACTGTGTGTCCTCTTTACAGCTATCAAGTCTCTGAGATGACCCAACTTGCAGAAC	2003
Qy	841	AAAGGAATTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTGACTATACCTCC	900
Db	2004	AAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTGACTATACCTCC	2063
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATGCTCTCAAGGCC	960
Db	2064	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATGCTCTCAAGGCC	2123
Qy	961	TCATGATCTACAGCCTGGAAAGAAATATCTCTCTGTGTGTTCAATATGGTGGTCTCTCA	1020
Db	2124	TCATGATCTACAGCCTGGAAAGAAATATCTCTCTGTGTGTTCAATATGGTGGTCTCTCA	2179
Qy	1021	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGA	1080
Db	2180	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGA	2239
Qy	1081	ACGTTATATGGGTCAACCTTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCAT	1140
Db	2240	ACGTTATATGGGTCAACCTTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCAT	2299
Qy	1141	GCAAGCAGAAAAGTTCCCTCTCTGAAACCAATCGTTTACTGTCTCTTACATGTTTCTCTGGA	1200

```
Db 2300 GCAAGCAGAAAGTTCCTCTGACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGA 2359
Qy 1201 TGAGATGTCCTATTTGACATACACAGTATATTAAGTATGTTTACTGAGGCTGGAA 1260
Db 2360 TGAGATGTCCTATTTGACATACACAGTATATTAAGTATGTTTACTGAGGCTGGAA 2419
Qy 1261 GCCATATGATTTACAGATCTATCTCCAGGAGAGACACAGCATAAGAGTTCTCGAATCGGG 1320
Db 2420 GCCATATGATTTACAGATCTATCTCCAGGAGAGACACAGCATAAGAGTTCTCGAATCGGG 2479
Qy 1321 AGAACATATGAACTGCACTTTTTCGACACTACCTTCAAGAAAACTTGGATCAGATATGC 1380
Db 2480 AGAACATATGAACTGCACTTTTTCGACACTACCTTCAAGAAAACTTGGATCAGATATGC 2539
Qy 1381 TGCTCTAAAGTGATATATTAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440
Db 2540 TGCTCTAAAGTGATATATTAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2599
Qy 1441 AACCAAATGAGGAGGTTTAATCAACAGAAAAACACAGAAATGATCATCACATTTTGATACC 1500
Db 2600 AACCAAATGAGGAGGTTTAATCAACAGAAAAACACAGAAATGATCATCACATTTTGATACC 2659
Qy 1501 TGCCATGTAACTCTACTCTGAAATATAATGTTGGTGCATGACAGGGGTCTACGGTTTGT 1560
Db 2660 TGCCATGTAACTCTACTCTGAAATATAATGTTGGTGCATGACAGGGGTCTACGGTTTGT 2719
Qy 1561 GGATGTAATCTTAATACCTTTAAACCCACATGCTCAAAATCAATGATACATATTCCTGAGA 1620
Db 2720 GGATGTAATCTTAATACCTTTAAACCCACATGCTCAAAATCAATGATACATATTCCTGAGA 2779
Qy 1621 GACCCAGCAATACCATAGATTAATCTAAAGAAAAAAGAAAAAAGAAAAA 1668
Db 2780 GACCCAGCAATACCATAGATTAATCTAAAGAAAAAAGAAAAAAGAAAAA 2827

RESULT 7
AX608731
LOCUS AX608731 4523 bp DNA linear PAT 17-FEB-2003
DEFINITION Sequence 8 from Patent WO0231134.
ACCESSION AX608731
VERSION AX608731.1 GI:28404301
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
AUTHORS Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
TITLE Novel serine protease genes related to dppiv
JOURNAL Patent: WO 0231134-A 8 18-APR-2002;
Ferring BV (NL)
FEATURES
    source
        1..4523
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match 98.8%; Score 1648.4; DB 6; Length 4523;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1663; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

Qy 1 AACAGTACAGCAATCCTTAAGTCACCTTTTAAAGATGTCAGAAATTAATGATGATGCTGA 60
Db 1164 AACAGTACAGCAATCCTTAAGTCACCTTTTAAAGATGTCAGAAATTAATGATGATGCTGA 1223
Qy 61 AGGAGGATCATAGATGTCATAGATAGGAACTTAATTCACCTTTTGAAGATCTATTGGA 120
Db 1224 AGGAGGATCATAGATGTCATAGATAGGAACTTAATTCACCTTTTGAAGATCTATTGGA 1283
Qy 121 AGGAGTGAATATATTGCCAGAGCTGGATGCACTCTCGAGGAAAAATATGCTTGGTCCAT 180
```

```
Db 1284 AGGAGTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGAAAAATATGCTTGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCAGACTCGCCTACAGTAGTGTGATCTCACCTGAATATTATAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGCCTACAGTAGTGTGATCTCACCTGAATATTATAT 1403
Qy 241 CCCAGTAGAAGATGATGTTATGAAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTCGT 300
Db 1404 CCCAGTAGAAGATGATGTTATGAAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTCGT 1463
Qy 301 GAGCCCACTAATTTATCTATGAAAGAAACAAACAGACACTCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCCACTAATTTATCTATGAAAGAAACAAACAGACACTCTGGATAAATATCCATGACATCTT 1523
Qy 361 TCATGTTTTTCCCAAGTTCACGAAAGAGGAAATTCAGTTTATTTTGGCTCTGAAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTTCACGAAAGAGGAAATTCAGTTTATTTTGGCTCTGAAATGCAA 1583
Qy 421 AACAGGTTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAACG 480
Db 1584 AACAGGTTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAACG 1643
Qy 481 ATCCAGTGTGGGCTCGCTCTCCAGTGATTTCAAGTGTCTTCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGTGGGCTCGCTCTCCAGTGATTTCAAGTGTCTTCAAGAGGAGATAGC 1703
Qy 541 AATTACCAGTGTGTAATGGAAAGTTCTTGGCCGCGCATGGATCTAATATCCAAAGTTGATGA 600
Db 1704 AATTACCAGTGTGTAATGGAAAGTTCTTGGCCGCGCATGGATCTAATATCCAAAGTTGATGA 1763
Qy 601 AGTCAGAGGCTGTGATATTTTGAAGGACACAAAGACTCCCTTTAGAGATCACTCTGTA 660
Db 1764 AGTCAGAGGCTGTGATATTTTGAAGGACACAAAGACTCCCTTTAGAGATCACTCTGTA 1823
Qy 661 CGTAGTCAGTACGTAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACA 720
Db 1824 CGTAGTCAGTACGTAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACA 1883
Qy 721 TTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTACCAAGAGAA 780
Db 1884 TTCTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTACCAAGAGAA 1943
Qy 781 TCACACTGTGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAAC 840
Db 1944 TCACACTGTGTGTCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAAC 2003
Qy 841 AAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTCTGACTATACTCTCC 900
Db 2004 AAAGGAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCCTCTCTGACTATACTCTCC 2063
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACATTTGTATGGGATGCTCTACAAGCC 960
Db 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACATTTGTATGGGATGCTCTACAAGCC 2123
Qy 961 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCTCA 1020
Db 2124 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGGTCTCTCA 2179
Qy 1021 GTTTGTCTATTGCTGGGGCCCCAGTCACCTCTGTGATCTTCTATCATACAGGATACACGGA 1080
Db 2180 GTTTGTCTATTGCTGGGGCCCCAGTCACCTCTGTGATCTTCTATCATACAGGATACACGGA 2239
Qy 1081 ACCTTATATGGGTCACTCCCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCAT 1140
Db 2240 ACCTTATATGGGTCACTCCCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCAT 2299
Qy 1141 GCAAGCAAGATGTTCCCTCTGACCAAAATCGCTTTACTGCTCTACTATCGTTTCCCTGGA 1200
Db 2300 GCAAGCAAGATGTTCCCTCTGACCAAAATCGCTTTACTGCTCTACTATCGTTTCCCTGGA 2359
Qy 1201 TGAGATGTCATTTTGGCACTACCAAGTATATTACTGAGTTTATTTTGTAGGGGCTGAAA 1260
```



Db	2360	TGGAATGTCCTATTTGGCACATACAGTATATTACTAGATTTTTTTAGTGAGGGCTGAAA	2419
Qy	1261	GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTGATCGGG	1320
Db	2420	GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTGATCGGG	2479
Qy	1321	AGACATATGAACTGATCTTTTGGCACTACCTTCAAGAAAACTTGGATCAGTATTGC	1380
Db	2480	AGACATATGAACTGATCTTTTGGCACTACCTTCAAGAAAACTTGGATCAGTATTGC	2539
Qy	1381	TGCTCTAAAGTGCATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	1440
Db	2540	TGCTCTAAAGTGCATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	2599
Qy	1441	AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATTTTGATACC	1500
Db	2600	AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATTTTGATACC	2659
Qy	1501	TGCATGTAACTACTCTCTGAAATAAATGTTGGTGCATGCGAGGGTCTACGGTTTGT	1560
Db	2660	TGCATGTAACTACTCTCTGAAATAAATGTTGGTGCATGCGAGGGTCTACGGTTTGT	2719
Qy	1561	GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATGATCATATTTCTTGAGA	1620
Db	2720	GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATGATCATATTTCTTGAGA	2779
Qy	1621	GACCCAGCAATACCATAGAAATTTACTAAAAAATAAAAAAAAAAAAAA 1668	
Db	2780	GACCCAGCAATACCATAGAAATTTACTAAAAAATAAAAAAAAAAAAAA 2827	
RESULT 8			
LOCUS	BC030688	3125 bp mRNA linear PRI 28-JUL-2005	
DEFINITION	Homo sapiens dipeptidyl peptidase 8, transcript variant 4, mRNA		
ACCESSION	BC030688		
VERSION	BC030688.2	GI:34190028	
KEYWORDS	MGC		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 3125)		
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S.P., Ziesberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, B., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.O., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krywinski, M.I., Skalski, U., Skalski, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.		
CONSTRM	Mammalian Gene Collection Program Team		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
REFERENCE	12477932		
AUTHORS	2 (bases 1 to 3125)		
CONSTRM	NIH MGC Project		
TITLE	Direct Submission		
JOURNAL	Submitted (24-MAY-2002) National Institutes of Health, Mammalian		

REMARK	Gene Collection (MGC), Bethesda, MD 20892-2590, USA
COMMENT	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>
	On Aug 25, 2003 this sequence version replaced gi:21265132.
	Contact: MGC help desk
	Email: <a href="mailto:cgapsb@mail.nih.gov">cgapsb@mail.nih.gov</a>
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
	cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA Sequencing by: Institute for Systems Biology
	<a href="http://www.systemsbio.org">http://www.systemsbio.org</a>
	contact: <a href="mailto:amad@systemsbiology.org">amad@systemsbiology.org</a>
	Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>
	Series: IRAC Plate; 33 Row; d Column; 5
	This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 37577092.
FEATURES	Location/Qualifiers
source	1..3125
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="MGC:26191 IMAGE:4822550"
	/tissue_type="Testis"
	/clone_lib="NIH MGC 97"
	/lab_host="DH10B"
	/note="Vector: pBluescriptR"
	1..3125
gene	/gene="DPP8"
	/notes="synonyms: FLJ20283, FLJ14920, MGC26191, MSTP141"
	/db_xref="GeneID:54878"
	/db_xref="MIM:606819"
CDS	319..2862
	/gene="DPP8"
	/codon_start=1
	/product="dipeptidyl peptidase 8, isoform 4"
	/protein_id="AAH30688.3"
	/db_xref="GI:54114905"
	/db_xref="GeneID:54878"
	/db_xref="MIM:606819"
	/translation="MWKSEQMKIKSGKNMAAAMETQLGVETPETAACENIESQD RPKLEPFYERYSWSQLKLADTKYGYMMAPHPDFVKNRDPDPSHRIYLL AMGENRENTLFYSEPKTINRAAALMSWKLDFLQATLDYGYSEELRERKR IGTVGASDYDHQSGTFLQAGSGIYHVDGPGQGTQPLRPNLVETSPNTRMDP KLCAPDPAIATHSNDIWSNIVFREERLTYYVNELANMEEDARSAGVATFVLOEE FDRYGYWCWCAETTPSGGKILRLIYENDESEVEIHTVSPMLETTRADSPRYPKT GTANPKVTFMSEIMIDAGRIIDVIDKELIQPFELFEGVEYIARAGWTEGKYAWS ILLDRSQTRLQIVLISPELFIPEDDVMERQLIESVPSVTPLIIEETDWINIH DIPHVTPQSEHEIEIFASECKTGFRLYKITSLKESKYKSSGGLPAPSPDKCP1 KERIATSGEWEVLGRHGSNIQVDEVRRLVPEGTQDSPLSHLYVSVYVNPGBVRL TDRGYSHSCISQHCDFETSKYGNOKNPHCVSLYKSSPEDDPTCKTKEPWTALDSA GVLPTPEISFESTGTFTYNGMLYKPHDLQPKKYPTVLFYIGGPQVQVNNREK GPVYFRINTLASIGYVVVDNRKSGHRLKPEGAFKYMVAIAGAPVTLWIFDVTGY TERYMGHPDQNEQGYLGSVMAQKFPSEPNRLLLHFLDENVFAHTSILLSFLV RACKPYDLQIYQPERHSIRVPSGSEHYELHLLHYLQENLGSRIAALKVI"
ORIGIN	
	Query Match 90.2%; Score 1505; DB 8; Length 3125;
	Best Local Similarity 91.9%; Pred. No. 0;
	Matches 1662; Conservative 0; Mismatches 0; Indels 147; Gaps 1;
Qy	1 AACAGGTACAGCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGCTGA 60
Db	1317 AACAGGTACAGCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGCTGA 1376
Qy	61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGAGATTTGGA 120
Db	1377 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGAGATTTGGA 1436
Qy	121 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTCTGAGGAAAAATATGCTGTGTCAT 180



```
Db 1437 AGAGTTGTAATATTTGTCAGAGCTGGATGGACTCCTCAGGGAATAATATGCTGGTCCAT 1496
Qy 181 CCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGTGATCTCACTGAATATTAT 240
Db 1497 CCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGTGATCTCACTGAATATTAT 1556
Qy 241 CCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGGCTGATCTGT 300
Db 1557 CCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGGCTGATCTGT 1616
Qy 301 GAGCCCACTAAATATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1617 GAGCCCACTAAATATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 1676
Qy 361 TCATGTTTTTCCCAAGTCAAGAGAGGAATAGAGTTATTTTTCCTCTGAATGCAA 420
Db 1677 TCATGTTTTTCCCAAGTCAAGAGAGGAATAGAGTTATTTTTCCTCTGAATGCAA 1736
Qy 421 AACAGGTTTCCGTCATTTATCAAAATACATCTATTTTAAAGGAGAGCAAAATATAACG 480
Db 1737 AACAGGTTTCCGTCATTTATCAAAATACATCTATTTTAAAGGAGAGCAAAATATAACG 1796
Qy 481 ATCCAGTGGTGGCTGCTCCAGTGCATTCAGTGCCTATCAAGAGAGATAGC 540
Db 1797 ATCCAGTGGTGGCTGCTCCAGTGCATTCAGTGCCTATCAAGAGAGATAGC 1856
Qy 541 AATTACAGTGGTGAATGGGAAGTTCCTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 600
Db 1857 AATTACAGTGGTGAATGGGAAGTTCCTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 1916
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACTGTGA 660
Db 1917 AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACTGTGA 1976
Qy 661 CGTAGTCAGTTACGTAAATCCTGGAGAGTGACAAAGGCTGACACCGTGGCTACTCACA 720
Db 1977 CGTAGTCAGTTACGTAAATCCTGGAGAGTGACAAAGGCTGACACCGTGGCTACTCACA 2036
Qy 721 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 780
Db 2037 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 2096
Qy 781 TCCACACTGTGTCTCCCTTTACAAGCTATCAAGTCTCAAGTATGACCCAACTTGCAGAAC 840
Db 2097 TCCACACTGTGTCTCCCTTTACAAGCTATCAAGTCTCAAGTATGACCCAACTTGCAGAAC 2156
Qy 841 AAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCC 900
Db 2157 AAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCC 2216
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCC 960
Db 2217 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCC 2276
Qy 961 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGTGTTTATATATGCGTGGTCTCA 1020
Db 2277 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGTGTTTATATATGCGTGGTCTCA 2336
Qy 1021 ----- 1020
Db 2337 GGTCAGTGGTGAATAATCGGTTTAAAGAGTCAAAGTATTTCCGTTGAATACCCTAGC 2396
Qy 1021 ----- 1020
Db 2397 CTCTCTAGTATTGTGGTTGTAGTATAGACAACAGGGGATCTGTGCACGAGGGCTTAA 2456
Qy 1021 -----GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTG 1053
Db 2457 ATTTGAAGCGCCTTTAAATATAAAATGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTG 2516
Qy 1054 GATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACCCAGAGAAATGAAC 1113
```

```
Db 2517 GATCTTCTATGATACAGGATACACGGAACTTTATATGGGTCAACCCTGACAGAAATGAACA 2576
Qy 1114 GGGCTATTACTTAGTATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCG 1173
Db 2577 GGGCTATTACTTAGTATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCG 2636
Qy 1174 TTTACTGCTCTTACATGGGTTTCTGTGATGAGAAATGCTCCATTTTGGACATACCAAGTATAT 1233
Db 2637 TTTACTGCTCTTACATGGGTTTCTGTGATGAGAAATGCTCCATTTTGGACATACCAAGTATAT 2696
Qy 1234 ACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAG 1293
Db 2697 ACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAG 2756
Qy 1294 ACACAGCATAAAGAGTTTCTCGAATCGGGAGAAACATTATGAATGCACTCTTTTGCACCT 1353
Db 2757 ACACAGCATAAAGAGTTTCTCGAATCGGGAGAAACATTATGAATGCACTCTTTTGCACCT 2816
Qy 1354 TCAGAAACCTTTGGATCAGTATTTGCTCTCTAAAGTGATATAATTTTGCCTGTGTA 1413
Db 2817 TCAGAAACCTTTGGATCAGTATTTGCTCTCTAAAGTGATATAATTTTGCCTGTGTA 2876
Qy 1414 GAACTCTCTGGTATACACTGGCTATTTACCAATGAGAGGTTTAAATCAACAGAAACA 1473
Db 2877 GAACTCTCTGGTATACACTGGCTATTTTACCAATGAGAGGTTTAAATCAACAGAAACA 2936
Qy 1474 CAGAAATGATCATCAGATTTTGAATACCTGCCATGTAACATCTAATCTCTGAAAAATAAATGT 1533
Db 2937 CAGAAATGATCATCAGATTTTGAATACCTGCCATGTAACATCTAATCTCTGAAAAATAAATGT 2996
Qy 1534 GGTGCCATGAGGGGCTACGGTTTGGTGTAGTAATCTAATACCTTAAACCCCATGCTC 1593
Db 2997 GGTGCCATGAGGGGCTACGGTTTGGTGTAGTAATCTAATACCTTAAACCCCATGCTC 3056
Qy 1594 AAAATCAATGATACATATTTCTGTGAGAGAGCCAGCAATACCAATAGAAATTAATAAAAA 1653
Db 3057 AAAATCAATGATACATATTTCTGTGAGAGAGCCAGCAATACCAATAGAAATTAATAAAAA 3116
Qy 1654 AAAAAAAAA 1662
Db 3117 AAAAAAAAA 3125

RESULT 9
AR631287 LOCUS AR631287 4676 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 20 from patent US 6844180.
ACCESSION AR631287
VERSION AR631287.1 GI:59770934
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4676)
AUTHORS Oi,S.; Akineanya,K.O.; Riviere,P.J.M. and Junien,J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 20 18-JAN-2005;
Ferring BV;;
NLX;

FEATURES
source Location/Qualifiers
1..4676
/mol_type="genomic DNA"

ORIGIN

Query Match 89.9%; Score 1501; DB 6; Length 4676;
Best Local Similarity 91.5%; Pred.No.0;
Matches 1663; Conservative 0; Mismatches 5; Indels 149; Gaps 1;

Qy 1 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATGTCAAGAAATTAATGATGCTGTA 60
Db 1164 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATGTCAAGAAATTAATGATGCTGTA 1223
```

QY	61	AGGAGGATCATAGATGTCATAGATAAGGAACTAATTAACCTTTTGAGATTCATTTGA	120
Db	1224	AGGAGGATCATAGATGTCATAGATAAGGAACTAATTAACCTTTTGAGATTCATTTGA	1283
QY	121	AGGAGTTGAATATATGCCAGAGCTGATGGACTCTGAGGGGAAATATGCTTGTCAT	180
Db	1284	AGGAGTTGAATATATGCCAGAGCTGATGGACTCTGAGGGGAAATATGCTTGTCAT	1343
QY	181	CCTACTAGATGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATATTTAT	240
Db	1344	CCTACTAGATGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATATTTAT	1403
QY	241	CCCAGTAGAGATGATGTTATGGAAGGACAGACATCAATGAGTCAGTGCCTGATCTGT	300
Db	1404	CCCAGTAGAGATGATGTTATGGAAGGACAGACATCAATGAGTCAGTGCCTGATCTGT	1463
QY	301	GAGCCCACTAATTAATCTATCAAGAAACAAGACATCTGGATAAATATCCATGACATCTT	360
Db	1464	GAGCCCACTAATTAATCTATCAAGAAACAAGACATCTGGATAAATATCCATGACATCTT	1523
QY	361	TCATGTTTTTCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGGCTCTGATGCAA	420
Db	1524	TCATGTTTTTCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGGCTCTGATGCAA	1583
QY	421	AACAGGTTTCGTCATTTATACAAATTAATCAATATTTTAAAGGAAACAAATATAACG	480
Db	1584	AACAGGTTTCGTCATTTATACAAATTAATCAATATTTTAAAGGAAACAAATATAACG	1643
QY	481	ATCCAGTGGTGGCTGCTCCAAAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGC	540
Db	1644	ATCCAGTGGTGGCTGCTCCAAAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGC	1703
QY	541	AAATACAGTGGTGAATGGGAAGTCTTGGCGGATGGATCTAATATCCAAAGTTGATGA	600
Db	1704	AAATACAGTGGTGAATGGGAAGTCTTGGCGGATGGATCTAATATCCAAAGTTGATGA	1763
QY	601	AGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGCTCCCTTTAGAGCATCACTGTA	660
Db	1764	AGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGCTCCCTTTAGAGCATCACTGTA	1823
QY	661	CGTAGTCAGTTAGTAACTCGAGAGGTGACAGGCTGACGCTGCTACTCACA	720
Db	1824	CGTAGTCAGTTAGTAACTCGAGAGGTGACAGGCTGACGCTGCTACTCACA	1883
QY	721	TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTATAGTATAGTACCAAGAA	780
Db	1884	TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTATAGTATAGTACCAAGAA	1943
QY	781	TCCACACTGTGTGCTCTTTTACAGCTATCAAGTCTGGAAGATGACCCAACTTCGAAAC	840
Db	1944	TCCACACTGTGTGCTCTTTTACAGCTATCAAGTCTGGAAGATGACCCAACTTCGAAAC	2003
QY	841	AAAGGAATTTTGGGCCACCAATTTTGAATTCAGAGGTCTCTTCTGACTATATCTCTCC	900
Db	2004	AAAGGAATTTTGGGCCACCAATTTTGAATTCAGAGGTCTCTTCTGACTATATCTCTCC	2063
QY	901	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAGCC	960
Db	2064	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAGCC	2123
QY	961	TCATGATCTACAGCTGGAAGAAATATCTCTGCTGCTGTTTCATATATGGTGTGTC----	1016
Db	2124	TCATGATCTACAGCTGGAAGAAATATCTCTGCTGCTGTTTCATATATGGTGTGTCGGGT	2183
QY	1017	-----	1016
Db	2184	CAATAGAAATTCAGCATCAGTGGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTC	2243
QY	1017	-----	1016
Db	2244	ATTGACTTAGATCGTGTGGGCATCCAGGCTGGTCTCTATGGAGGATACCTCTCCCTGATG	2303
QY	1017	-----CTCAGTTGCTATTGCTGGGGCCCCAGTCACCTGT	1051

Db	2304	GCATTAATGTCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCATCTG	2363
QY	1052	TGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTCACCAGAAATGAA	1111
Db	2364	TGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTCACCAGAAATGAA	2423
QY	1112	CAGGCTATTAATCTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAT	1171
Db	2424	CAGGCTATTAATCTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAT	2483
QY	1172	CGTTTACTGCTCTTACATGTTTCTGGATGAGAAATGTCATTTTGCACATACCAGTATA	1231
Db	2484	CGTTTACTGCTCTTACATGTTTCTGGATGAGAAATGTCATTTTGCACATACCAGTATA	2543
QY	1232	TTACTGAGTGTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAG	1291
Db	2544	TTACTGAGTGTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAG	2603
QY	1292	AGACACAGCATPAAGAGTTCTGAAATCGGAGAAACATTATGAACCTGTCATCTTTGCACTAC	1351
Db	2604	AGACACAGCATPAAGAGTTCTGAAATCGGAGAAACATTATGAACCTGTCATCTTTGCACTAC	2663
QY	1352	CTTCAAGAAAACCTTGGATCAGCTATGCTGCTCTAAAAGTGATATAATTTTGACCTGTG	1411
Db	2664	CTTCAAGAAAACCTTGGATCAGCTATGCTGCTCTAAAAGTGATATAATTTTGACCTGTG	2723
QY	1412	TAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAA	1471
Db	2724	TAGAACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAA	2783
QY	1472	CACAGAAATGATCATCATTTTGTATACCTGTCATGTAACATCTCTCTGAAAAATAAT	1531
Db	2784	CACAGAAATGATCATCATTTTGTATACCTGTCATGTAACATCTCTCTGAAAAATAAT	2843
QY	1532	GTGGTGCATCAGGAGGTCTACGGTTTGGTGTAGTATCTAATACCTTAACCCCATGTC	1591
Db	2844	GTGGTGCATCAGGAGGTCTACGGTTTGGTGTAGTATCTAATACCTTAACCCCATGTC	2903
QY	1592	TCAAAATCAATGATACATATCTCTGAGAGCCAGCAATACCATAGAAATTTACTAAAA	1651
Db	2904	TCAAAATCAATGATACATATCTCTGAGAGCCAGCAATACCATAGAAATTTACTAAAA	2963
QY	1652	AAAAAAAAAAAAAAAAAAAA 1668	
Db	2964	AAAAAAAAAAAAAAAAAAAA 2980	
RESULT 10			
AX608743			
LOCUS			
DEFINITION			
AX608743			
ACCESSION			
AX608743.1			
VERSION			
AX608743.1			
KEYWORDS			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
1			
AUTHORS			
Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.			
TITLE			
Novel serine protease genes related to dppiv			
JOURNAL			
Patent: WO 0231134-A 20 18-APR-2002;			
Ferring BV (NL)			
FEATURES			
source			
1..4676			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
LOCATION/Qualifiers			
1..4676			
ORIGIN			
Query Match			
89.9%; Score 1501; DB 6; Length 4676;			





RESULT 12	AX878058	Sequence 12963 from Patent EP1074617.	2161 bp	DNA	linear	PAT 17-DEC-2003
LOCUS	AX878058					
DEFINITION	AX878058					
ACCESSION	AX878058					
VERSION	AX878058.1	GI:40032794				
KEYWORDS						
SOURCE						
ORGANISM						
	Homo sapiens (human)					
	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;					
	Hominidae; Homo.					
REFERENCE						
AUTHORS	Ota, T., Isegai, T., Nishikawa, T., Hayaishi, K., Saito, K., Yamamoto, J., Ieshi, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuka, T.					
TITLE	Primers for synthesising full-length cDNA and their use					
JOURNAL	Patent: EP 1074617-A 12963 07-FEB-2001;					
	Research Association for Biotechnology					
FEATURES						
source	1..2161	Location/Qualifiers				
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
CDS	16..1914	note="unnamed protein product"				
	/codon_start=1					
	/protein_id="CA889991.1"					
	/db_xref="GI:40032795"					
	/translations="MOPKLCFADPDWIAFIHSNDIWSIVTREERLTIVVHLEANNH BEDARSAGVATFVLQEFDRYSQVWMCPEATTPSGKILRLILVEENDESEVERIHVT SPMLERTRADSFVRYKPTGILDRSPKVTFMSEIMIDAEGRIDVIDKELIQPEILPFGV EYIARLEATGPEGKVAWSILLDRSOTRLQVILSPLELTPVEDDQWESQRIEIVPDSV TPLIIVBETDIDINHTDILPHVPPQSHREIBFIPASCKTGPHRLVKITSLIKESKY HHLVSVVNGPEVTALDRPSVSHSCISQHCDFRISKYKNQKPHCVSLYKLSSPED DCTCKTEPWTILDSAGLPDYPTEPEIFESPESFTGFLYGLMLKPHLDLPQKKRPIV LFIYGGQVOLVNNRPGVKYPRLLATSLGTVVVINDRGSCHRGKLPFGKAFKYMV AIAQAPVTALFVDTGYTRYKMGHPDNEQGYLGSSVAMQAKFPFSEPNRLLLHGLFL DENVHPAHTSILLFLVRAKGVLDLIQIPQENHSIRVPSGEHYELHLLHYLQENLGS RIALKRKVI"					
ORIGIN						
Query Match	89.1%;	Score 1487.4;	DB 6;	Length 2161;		
Best Local Similarity	91.7%;	Prod. No. 0;				
Matches 1645;	Conservative	0;	Mismatches 1;	Indels 147;	Gaps 1;	
QY	1	AACAGGTACAGCAAACTCCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGA	60			
Db	369	AACAGGTACAGCAAACTCCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGA	428			
QY	61	AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTGA	120			
Db	429	AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTGA	488			
QY	121	AGGAGTTGAATATATTCGCAGAGCTCGATGGACTCCTCAGGGAATAATGCTGGTCCAT	180			
Db	489	AGGAGTTGAATATATTCGCAGAGCTCGATGGACTCCTCAGGGAATAATGCTGGTCCAT	548			
QY	181	CCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGTGATCTCACTGAAATATTAT	240			
Db	549	CCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGTGATCTCACTGAAATATTAT	608			
QY	241	CCAGTGAAGATGATGTTATGAAAGGAGAGACTCATTCAGTCAGTGCCTGATCTGT	300			
Db	609	CCAGTGAAGATGATGTTATGAAAGGAGAGACTCATTCAGTCAGTGCCTGATCTGT	668			
QY	301	GACGCCACTAATTATCTATGATAAACAACAGACATCTGGATAAATATCCATGACATCTT	360			
Db	669	GACGCCACTAATTATCTATGATAAACAACAGACATCTGGATAAATATCCATGACATCTT	728			
QY	361	TCATGTTTTTCCCCAAAGTCACGAAGGAAATTCAGTTTATTTTTCCTCTCAATGCAAC	420			

QY 1354 TCAAGAAAACCTTGGATCAGGATTTGCTGCTCTAAAGTGATATAAATTTTGACCTGTGTA 1413  
|||||  
Db 1869 TCAAGAAAACCTTGGATCAGGATTTGCTGCTCTAAAGTGATATAAATTTTGACCTGTGTA 1928  
|||||  
QY 1414 GAATCTCTGGTATACACATGCGGTATTAACCAATGAGGAGTTTAAATCAACAGAAAACA 1473  
|||||  
Db 1929 GAATCTCTGGTATACACATGCGGTATTAACCAATGAGGAGTTTAAATCAACAGAAAACA 1988  
|||||  
QY 1474 CAGAAATGATCATCAATTTTGATACCTGCGCATGTAACATCTACTCTCGAAATAAATGT 1533  
|||||  
Db 1989 CAGAAATGATCATCAATTTTGATACCTGCGCATGTAACATCTACTCTCGAAATAAATGT 2048  
|||||  
QY 1534 GGTGCCATCAGGGGTCTACGGTTTGTGGTAGTAATCTAATCAACCCCATGCTC 1593  
|||||  
Db 2049 GGTGCCATCAGGGGTCTACGGTTTGTGGTAGTAATCTAATCAACCCCATGCTC 2108  
|||||  
QY 1594 AATCAATCAATGATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTAAT 1646  
|||||  
Db 2109 AATCAATCAATGATACATATTTCTGAGAGACCCAGCAATACCATAGAAATTAAT 2161  
|||||

RESULT 13  
AK027826 2161 bp mRNA linear PRI 30-JAN-2004  
LOCUS Homo sapiens cDNA FLJ14920 fis, clone PLACE1007416, weakly similar  
DEFINITION to DIPEPTIDYL PEPTIDASE IV (EC 3.4.14.5).  
ACCESSION AK027826  
VERSION AK027826.1 GI:14042789  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1  
Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,  
Wakabatu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,  
Sekihe, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,  
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,  
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,  
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,  
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,  
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,  
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,  
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,  
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,  
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,  
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,  
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,  
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,  
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,  
Terahima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,  
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,  
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,  
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,  
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,  
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,  
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,  
Noguchi, S., Itoh, T., Shigetake, K., Senba, T., Matsumura, K.,  
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,  
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,  
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,  
Nagata, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,  
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.  
Complete sequencing and characterization of 21,243 full-length  
human cDNAs  
Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL  
PUBMED 14702039  
REFERENCE 2  
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,  
Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

FEATURES  
source

CDS

ORIGIN

Query Match 89.1%; Score 1487.4; DB 8; Length 2161;  
Best Local Similarity 91.74; Pred. No. 0;  
Matches 1645; Conservative 0; Mismatches 1; Indels 147; Gaps 1;  
QY 1 AACAGGTACAGCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGA 60  
Db 369 AACAGGTACAGCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGA 428  
QY 61 AGGAAGGATCATAGATGTCATAGATAGGAACCTTAATCAACCTTTTGAGATTCCTATTGA 120  
Db 429 AGGAAGGATCATAGATGTCATAGATAGGAACCTTAATCAACCTTTTGAGATTCCTATTGA 488  
QY 121 AGGAGTTGAATATATGCCAGAGCTGGACTCTGAGGAGAAATAATGCTTGCTCCAT 180  
Db 489 AGGAGTTGAATATATGCCAGAGCTGGACTCTGAGGAGAAATAATGCTTGCTCCAT 548  
QY 181 CCTACTAGATCGCTCCAGACTCGCTCTAGATAGTGTGTGATCTCACCTGGAATATTTAT 240  
Db 549 CCTACTAGATCGCTCCAGACTCGCTCTAGATAGTGTGTGATCTCACCTGGAATATTTAT 608  
QY 241 CCCAGTAGAAGATGATGTTATGGAAAGGCGAGAGACTCATTAGTCTAGTCCCTGATTCGT 300  
Db 609 CCCAGTAGAAGATGATGTTATGGAAAGGCGAGAGACTCATTAGTCTAGTCCCTGATTCGT 668  
QY 301 GACCCCACTAATATCTATCAAGAAACACAGACATCTCGATAAATATCATGACATCTT 360  
|||||

Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,  
Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,  
Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,  
Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.  
NEDO human cDNA sequencing project  
Unpublished  
3 (bases 1 to 2161)  
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing;  
Research Association for Biotechnology; cDNA library construction;  
5'- & 3'-end one pass sequencing and clone selection; Helix  
Research Institute (supported by Japan Key Technology Center etc.)  
and Department of Virology, Institute of Medical Science,  
University of Tokyo.  
Location/Qualifiers  
1. .2161  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE1007416"  
/tissue\_type="placenta"  
/clone\_lib="PLACE1"  
/notes="cloning vector: pME18SFL3"  
16. .1914  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB55395.1"  
/db\_xref="GI:14042790"  
/translation="MDPKLCPADPDWIAPIHSNDIWNISNIVTREERTLYVHNELANN  
EEDARGAGVATFVLOEFEDRYGYWVCKPAETTPGGKILRLILEENDESEVEIITHVT  
SPMLTERRADSRFRYKPTGTANKPTFKMSEIMIDAGRI IDVIDKELIQPEILPEGV  
EYATARGWTPEGYKWSILLDSQRLQIVLISPELFIPEVDDWNERQRLIESVPSDV  
TPILIIYEETDWINIHDFHFPQSBEHEIEFIIPASECKTGFRHLKYITILKESKY  
KRLLGGLPAPDFKPIKEEIAITPGSEVEELRGHSNIQVDEVRLLYVFEFGKQSPLE  
HHLYVSVYVNGVEVRLTRDPSYSHSCISQHCDFIISKYSNQKPHCVSLYKJASPED  
DPYCKTKEFWATILDSAGLPDYTPPEIFPSFSTGTFTLYGMLYKPHLOKFGKPYPTV  
LFYIGPQVLVNNREKGVYFRLATLASLGVVVVVDNRGSHRGLAFEGAFYKMY  
AIGAPVTLMIFDYTYTRYMGHPDQEQYGLSVAMQAEKFPSEPNRLLLHGLF  
DENVHAFTSILLSLFLVRAGKPYDQIYPPQERHSIRVPESGEHYELHLLHQLQENIGLS  
RIAAALKVI"



Db	669	GACGCCACTAATTCTATGAGAAACACACACATCTGGATAAATATCCATGACATCTT	728
Qy	361	TCATGTTTTTCCCAAGTCACGAGAGGAATTCAGTTTATTTTTCGCTCTGATGCAA	420
Db	729	TCATGTTTTTCCCAAGTCACGAGAGGAATTCAGTTTATTTTTCGCTCTGATGCAA	788
Qy	421	AACAGGTTCCGTCATTTATACAAATTCATCTATTTTAAAGGAAAGCAAAATATAACG	480
Db	789	AACAGGTTCCGTCATTTATACAAATTCATCTATTTTAAAGGAAAGCAAAATATAACG	848
Qy	481	ATCCAGTGTGGCTGCCCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAAAGAGGATAGC	540
Db	849	ATCCAGTGTGGCTGCCCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAAAGAGGATAGC	908
Qy	541	AATTACAGTGTGAATGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA	600
Db	909	AATTACAGTGTGAATGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA	968
Qy	601	AGTCAGAAGCGTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTA	660
Db	969	AGTCAGAAGCGTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTA	1028
Qy	661	CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTACTCACA	720
Db	1029	CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTACTCACA	1088
Qy	721	TTCTTCCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTAAATAGTAAACCAAGRA	780
Db	1089	TTCTTCCTGCATCAGTCAGCACTGTGACTTCTTTTATAGTAAATAGTAAACCAAGRA	1148
Qy	781	TCACACTGTGTGCTCCCTTTACAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAG	840
Db	1149	TCACACTGTGTGCTCCCTTTACAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAG	1208
Qy	841	AAAGGAATTTTGGGCCACATTTTGGATTCAGAGGTCCTCTCTCTGACTATACCTCTCC	900
Db	1209	AAAGGAATTTTGGGCCACATTTTGGATTCAGAGGTCCTCTCTCTGACTATACCTCTCC	1268
Qy	901	AGAAATTTTCTTTTGAAGTACTCTGGAATTTACATTTGATGGGATGCTCTCAAGCC	960
Db	1269	AGAAATTTTCTTTTGAAGTACTCTGGAATTTACATTTGATGGGATGCTCTCAAGCC	1328
Qy	961	TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGTCTTATATATATGGTGGTCTCA	1020
Db	1329	TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGTCTTATATATATGGTGGTCTCA	1388
Qy	1021	-----	1020
Db	1389	GGTGCAAGTGTGTAATCGTGTAAAGAGTCAAGTATTTTCOGCTTGAATACCCCTAGC	1448
Qy	1021	-----	1020
Db	1449	CTCTCTAGGTTATGTGGTTGTAGTATAGACAAACAGGGGATCCTGTCCAGGGGCTTAA	1508
Qy	1021	-----GGTTGCTATTGCTGGGGCCCAAGTCACTCTGTG	1053
Db	1509	ATTGNAAGGCGCTTTAATATAAATGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTG	1568
Qy	1054	GATCTCTATGATACAGGATACAGGAACGTTTATATGGTCAACCTGTGACCAAGATGAACA	1113
Db	1569	GATCTCTATGATACAGGATACAGGAACGTTTATATGGTCAACCTGTGACCAAGATGAACA	1628
Qy	1114	GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCG	1173
Db	1629	GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCG	1688
Qy	1174	TTTACTGCTTTACATGGTTTCTGGATGAATGTCCATTTTGGCATAACCATGATATT	1233
Db	1689	TTTACTGCTTTACATGGTTTCTGGATGAATGTCCATTTTGGCATAACCATGATATT	1748
Qy	1234	ACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCCTCAGAGAG	1293
Db	1749	ACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCCTCAGAGAG	1808

Qy	1294	ACACAGCATTAAGAGTTCTCTGAATCGGAGAGAAATTTATGAATCGATCTTTTGCATACCT	1353
Db	1809	ACACAGCATTAAGAGTTCTCTGAATCGGAGAGAAATTTATGAATCGATCTTTTGCATACCT	1868
Qy	1354	TCAAGAAAACTTTGGATCAGTATTCCTCTCTTAAAGTGATATAATTTTGACCTGTGTA	1413
Db	1869	TCAAGAAAACTTTGGATCAGTATTCCTCTCTTAAAGTGATATAATTTTGACCTGTGTA	1928
Qy	1414	GAACTCTCTGTTATACATCGGCTTATTAACCAAAATGAGGAGGTTTAAATCAACAGAAAAACA	1473
Db	1929	GAACTCTCTGTTATACATCGGCTTATTAACCAAAATGAGGAGGTTTAAATCAACAGAAAAACA	1988
Qy	1474	CAGAAATTCATCATCATATTTTGATACCTGCCATGTAAACATCTACTCTCGAAAAATAATGT	1533
Db	1989	CAGAAATTCATCATCATATTTTGATACCTGCCATGTAAACATCTACTCTCGAAAAATAATGT	2048
Qy	1534	GGTGCCATGACAGGGGCTCTAGGTTTGTGTAGTAAATCTAATACCTTAACCCACATGCTC	1593
Db	2049	GGTGCCATGACAGGGGCTCTAGGTTTGTGTAGTAAATCTAATACCTTAACCCACATGCTC	2108
Qy	1594	AAAATCAATGATACATATTTCTTGAGAGAGCCAGCAATACCAATAGAAATTAAT	1646
Db	2109	AAAATCAATGATACATATTTCTTGAGAGAGCCAGCAATACCAATAGAAATTAAT	2161

RESULT 14  
AF221634  
LOCUS AF221634  
DEFINITION Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, complete cds.  
ACCESSION AF221634  
VERSION AF221634.1 GI:11095187  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3127)  
AUTHORS Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.  
TITLE Cloning, expression and chromosomal localization of a novel human  
dipeptidyl peptidase (Dpp) IV homolog, DPP8  
JOURNAL Eur. J. Biochem. 267 (20), 6140-6150 (2000)  
PUBMED 11012666  
AUTHORS 2 (bases 1 to 3127)  
TITLE Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.  
JOURNAL Direct Submission  
PUBMED  
AUTHORS Submitted (06-JAN-2000) A.W. Moffrow Gastroenterology and Liver  
JOURNAL Centre, Centenary Institute of Cell Biology and Cancer Medicine,  
Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia  
FEATURES  
source  
1..3127  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q22"  
1..3127  
/gene="DPP8"  
214..2862  
/gene="DPP8"  
/note="serine protease; member of prolol oligopeptidase S9b  
family belonging to enzyme clan SC; alternatively spliced  
forms of this mRNA are found in AF221635, AF221636 and  
AF221637"  
/codon\_start=1  
/product="dipeptidyl peptidase 8"  
/protein\_id="AAG29766.1"  
/db\_xref="GI:11095188"  
/translation="MAAAETEQLGVFEFTADCEENIESODRPFKLEPFYVYRWSQ  
LKKLLADTKYHYNMKAPDFKRNDDPHSDRIYVLAMSGNRRNTLPYSEI  
PKTINRAAVLMLSWKPLDLFOATLDYGMYSREBELLRKRKTGTVGIAVDYHQSGG



TFLEPAGSGIYHVKGDPQFTQOPLRPNLVETSCPNIRMBPKLCPADPDWIAPIHNS  
DIWISNIVTREERLTYVHNELANNEEDARSAGVATFVLQEEFDRIYGYWCPKRAEIT  
PSGKIILIRIILENDESEVEI IHVTSPLMETRRADSFYPIKGTGTANPKVTFPMSEIMI  
DAEGRIIDIVLKLIOFFELFEGVEIYARAGWTEGKYAWSILLDSQRLQIWLIS  
PELFI PVEDDVMERQRLIESVDSVPLIIYEETTDIWINIHDIFHVPPOSHHEEIEF  
IPASECKTGFRLHYKITSILKESKYKSSGSLPAPSPDFKCPKBEIIAITSGEWEVLGR  
HGSNIQVDEVRILVYPEGKDSPLEHLHVSVNVPGEVTRBLTDRGYSHSCCISQHCD  
FFTSKYSNQKPHCVSLYKLSPPEDDPTCKTEFWATILDSAGPLDPYTPPEIIPSPBS  
VVVIDNRSGHRLKPEGAFKYMGGQIEIDDDQVEGLQYLASRYDFIDLDRVGIHGWY  
GGYLSMALMORLDI FRVLAIGAPVTLWIFDYGTGYRMGHGPDQNEGYVILGSVAMQ  
AEKFPSPBNLLHGHFLDENVHFAHTSILLSFLVRAGKPYDLQIYQBERHSIRVPS  
GEHYELHLHYLQENLGSRIAALKVI"

ORIGIN

Query Match 81.1%; Score 1354; DB 8; Length 3127;

Best Local Similarity 84.7%; Pred. No. 9.2e-296;

Matches 1664; Conservative 0; Mismatches 0; Indels 300; Gaps 1;

QY	1	AACAGGTACAGCAATCCTAAAGTCATCTTTAAGATGTCAGAAATAATGATGATGCTGA	60
DB	1164	AACAGGTACAGCAATCCTAAAGTCATCTTTAAGATGTCAGAAATAATGATGATGCTGA	1223
QY	61	AGGAAGATCATAGATGTCATAGATAGGAATCTAAATCAACCTTTTGAGATTCCTATTGA	120
DB	1224	AGGAAGATCATAGATGTCATAGATAGGAATCTAAATCAACCTTTTGAGATTCCTATTGA	1283
QY	121	AGGAGTTGAATATATTGCGAGAGCTGGATGGACTCTCTGAGGGAATAATGCTGGTCCAT	180
DB	1284	AGGAGTTGAATATATTGCGAGAGCTGGATGGACTCTGAGGGAATAATGCTGGTCCAT	1343
QY	181	CCTACTAGATGCTGCCAGATCCGCTACAGATAGTGTGATCTCACCTGAATATTAT	240
DB	1344	CCTACTAGATGCTGCCAGATCCGCTACAGATAGTGTGATCTCACCTGAATATTAT	1403
QY	241	CCCAGTAGAGATGATGTTATGGAAGGAGAGACTCATTTAGTCAGTGCCTGATCTGT	300
DB	1404	CCCAGTAGAGATGATGTTATGGAAGGAGAGACTCATTTAGTCAGTGCCTGATCTGT	1463
QY	301	GAGCCACTAATTTATCTATGAAGAAACAACAGACATCTGGATAATATCCATGACATCTT	360
DB	1464	GAGCCACTAATTTATCTATGAAGAAACAACAGACATCTGGATAATATCCATGACATCTT	1523
QY	361	TCATGTTTTTCCCAAGTCAGAGAGAAATGAGTTATTTTGTGCTCTGAATGCAA	420
DB	1524	TCATGTTTTTCCCAAGTCAGAGAGAAATGAGTTATTTTGTGCTCTGAATGCAA	1583
QY	421	AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAACCAATATAACG	480
DB	1584	AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAACCAATATAACG	1643
QY	481	ATCCAGTGGTGGCTGCTGCTCCAAAGTCATTTCAAGTGTCTCTATCAAGAGGAGATAGC	540
DB	1644	ATCCAGTGGTGGCTGCTGCTCCAAAGTCATTTCAAGTGTCTCTATCAAGAGGAGATAGC	1703
QY	541	AAATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	600
DB	1704	AAATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	1763
QY	601	AGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTTGAAGCATACCTGTGA	660
DB	1764	AGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTTGAAGCATACCTGTGA	1823
QY	661	CGTAGTCAGTTAGTAAATCCCTGGAGAGTGACAGGCTGACTGACCGTGGCTACTCACA	720
DB	1824	CGTAGTCAGTTAGTAAATCCCTGGAGAGTGACAGGCTGACTGACCGTGGCTACTCACA	1883
QY	721	TTCTTGCTGATCAGTCAGACTGTGACTTCTTTTAAAGTAAAGTATAGTACCAAGAGAA	780
DB	1884	TTCTTGCTGATCAGTCAGACTGTGACTTCTTTTAAAGTAAAGTATAGTACCAAGAGAA	1943
QY	781	TCACACTGTGTGCTCCTTTACAGCTATCAAGTCTCAAGTCTCGAAGATGACCCCACTTGC	840
DB		TCACACTGTGTGCTCCTTTACAGCTATCAAGTCTCGAAGTCTCGAAGTATATATATATAT	

DB	1944	TCACACTGTGTGTCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGC	2003
QY	841	AAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCC	900
DB	2004	AAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCC	2063
QY	901	AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATATGGGATGCTCTCAAGCC	960
DB	2064	AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATATGGGATGCTCTCAAGCC	2123
QY	961	TCATCATCTACAGCTGGGAAAGAAATATCTCTAGTGTCTCTCATATATGTTGGTCTCTCA	1020
DB	2124	TCATCATCTACAGCTGGGAAAGAAATATCTCTAGTGTCTCTCATATATGTTGGTCTCTCA	2183
QY	1021	-----	1020
DB	2184	GGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGC	2243
QY	1021	-----	1020
DB	2244	CTCTTAGTTATGTGTTGTAGTATAGACAAAGGGATCTCTGTCAACGAGGGCTTAA	2303
QY	1021	-----	1020
DB	2304	ATTTGAAGCGCCCTTTTAAATATATAAATGGGTCAAAATGACATGAGTGAAGG	2363
QY	1021	-----	1020
DB	2364	ACTCCATATCTAGCTTCTCGATATGATTTTATTGACTAGATCGTGTGGGCATCCACGG	2423
QY	1021	-----	1020
DB	2424	CTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTGAGATATCTTCA	2483
QY	1021	-----	1020
DB	1021	GGTGTCTTATGCTGGGCGCCAGTCACTCTGTGGATCTTATGATACAGGATACACGGA	1080
DB	2484	GGTGTCTTATGCTGGGCGCCAGTCACTCTGTGGATCTTATGATACAGGATACACGGA	2543
QY	1081	ACGTTATATGGGTCACTCTGACAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCAT	1140
DB	2544	ACGTTATATGGGTCACTCTGACAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCAT	2603
QY	1141	GCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTACTGTCTTTACATGGTTCTGGA	1200
DB	2604	GCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTACTGTCTTTACATGGTTCTGGA	2663
QY	1201	TGAGATGTCCATTTTGCACATACCATGATATTTACTGAGCTTTTATGACGGCTGGAA	1260
DB	2664	TGAGATGTCCATTTTGCACATACCATGATATTTACTGAGCTTTTATGACGGCTGGAA	2723
QY	1261	GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGAGTTCTTGAATCGGG	1320
DB	2724	GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGAGTTCTTGAATCGGG	2783
QY	1321	AGAACTATTAAGATGCTCTTTTGCACCTTCAAGAAAACTTTGGATCACTGATTTGC	1380
DB	2784	AGAACTATTAAGATGCTCTTTTGCACCTTCAAGAAAACTTTGGATCACTGATTTGC	2843
QY	1381	TGCTCTAAAGTATATAATTTTGAACCTGTGTAGAACTCTCTGTGTATACACTGGCTATT	1440
DB	2844	TGCTCTAAAGTATATAATTTTGAACCTGTGTAGAACTCTCTGTGTATACACTGGCTATT	2903
QY	1441	AACCAATAGGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGTATACC	1500
DB	2904	AACCAATAGGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGTATACC	2963
QY	1501	TGCCATGTAAACATCTACTCTCTGAAAAATAATGTGGTCCATGCAGGGGTCTACGGTTGT	1560
DB	2964	TGCCATGTAAACATCTACTCTCTGAAAAATAATGTGGTCCATGCAGGGGTCTACGGTTGT	3023
QY	1561	GGTAGTATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGA	1620
DB	3024	GGTAGTATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGA	3083



Qy 1081 ACGTTATATGGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140  
Db 2559 ACGTTATATGGGTCAACCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2618  
Qy 1141 GCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGA 1200  
Db 2619 GCACGAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGA 2678  
Qy 1201 TGAGAAATGTCATTTTGGACATACACAGTATATTTACTAGTTTCTTGTAGGGCTGGAAA 1260  
Db 2679 TGAGAAATGTCATTTTGGACATACACAGTATATTTACTAGTTTCTTGTAGGGCTGGAAA 2738  
Qy 1261 GCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAGCATAAAGAGTTCTTGAATCGGG 1320  
Db 2739 GCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAGCATAAAGAGTTCTTGAATCGGG 2798  
Qy 1321 AGAACATTTATGAATGCTATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCAGTATTGC 1380  
Db 2799 AGAACATTTATGAATGCTATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCAGTATTGC 2858  
Qy 1381 TGCTCTAAAGTGATATTAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440  
Db 2859 TGCTCTAAAGTGATATTAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2918  
Qy 1441 AACCRAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCACATTTTGATACC 1500  
Db 2919 AACCRAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTGATCATCACATTTTGATACC 2978  
Qy 1501 TGCCATGTAACTCTACTCTGAAATAAATGTGGTGCCATGCGGGGTCTACGGTTTGT 1560  
Db 2979 TGCCATGTAACTCTACTCTGAAATAAATGTGGTGCCATGCGGGGTCTACGGTTTGT 3038  
Qy 1561 GGTAGTAATCTAATACCTTTAACCACATGCTCAAAATCAAATGATACATATTTCTCTGAGA 1620  
Db 3039 GGTAGTAATCTAATACCTTTAACCACATGCTCAAAATCAAATGATACATATTTCTCTGAGA 3098  
Qy 1621 GACCCAGCAATACCATTAAGAAATTTCTAAAAAATTTTTTTTTTTTTTTTTTTTTT 1665  
Db 3099 GACCCAGCAATACCATTAAGAAATTTCTAAAAAATTTTTTTTTTTTTTTTTTTTTT 3143

Search completed: May 4, 2006, 18:51:59  
Job time : 8334.8 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:50:44 ; Search time 939.948 Seconds  
(without alignments)  
11834.025 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 1669  
Sequence: 1 aacaggtacagcaatccta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues  
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq\_21.\*
- 1: Geneseqn1980s.\*
  - 2: Geneseqn1990s.\*
  - 3: Geneseqn2000s.\*
  - 4: Geneseqn2001as.\*
  - 5: Geneseqn2001bs.\*
  - 6: Geneseqn2002as.\*
  - 7: Geneseqn2002bs.\*
  - 8: Geneseqn2003as.\*
  - 9: Geneseqn2003bs.\*
  - 10: Geneseqn2003cs.\*
  - 11: Geneseqn2003ds.\*
  - 12: Geneseqn2004as.\*
  - 13: Geneseqn2004bs.\*
  - 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1669	100.0	1669	4 AAC85696	AAC85696 Nucleotid
2	1659	99.4	2668	6 ABN59775	ABN59775 Novel hum
3	1659	99.4	2842	6 ABN59774	ABN59774 Novel hum
4	1648.4	98.8	4523	6 ABK83325	ABK83325 cDNA enco
5	1614.4	96.7	2702	13 ADT04072	ADT04072 Human pro
6	1501	89.9	4676	6 ABK83331	ABK83331 cDNA enco
7	1493.4	89.5	2952	10 ACA92421	ACA92421 DNA enco
8	1487.4	89.1	2161	4 AAH15009	AAH15009 Human CDN
9	1399	83.8	2349	14 ADV43982	ADV43982 Human psy
10	1353.4	81.1	3143	6 AAH99934	AAH99934 cDNA enco
11	1350	80.9	2083	7 ADR41222	ADR41222 Human CD-
12	1349.6	80.9	4829	6 ABK83327	ABK83327 cDNA enco
13	1347	80.7	3120	4 AAC85694	AAC85694 Nucleotid
14	1347	80.7	3120	6 AAD38956	AAD38956 Human dip
15	1344	80.5	3106	6 ABK12892	ABK12892 Human pro
16	1343.6	80.5	4685	6 ABK83332	ABK83332 cDNA enco
17	1341	80.3	2929	10 ACA92425	ACA92425 DNA enco
18	1337	80.1	2797	12 ADL13374	ADL13374 Human ste
19	1266.8	75.9	2696	12 ADI16386	ADI16386 Human pro

20	1250.6	74.9	2510	6 AAD23843	AAD23843 Human pro
21	1209.4	72.5	1821	6 ABV76411	ABV76411 Dipeptidyl
22	1183.6	70.9	2463	10 ACA92424	ACA92424 DNA enco
23	1089	65.2	2649	14 ADV43981	ADV43981 Human psy
24	1085.8	65.1	2649	8 ABX12255	ABX12255 cDNA enco
25	1085.8	65.1	2671	6 ABK83322	ABK83322 cDNA enco
26	1084.4	65.0	2643	6 AAH99935	AAH99935 Coding se
27	1040.6	62.3	2251	10 ADE79035	ADE79035 Human pro
28	881	52.8	1197	4 AAC85695	AAC85695 Nucleotid
29	765.2	45.8	4309	6 ABK83328	ABK83328 cDNA enco
30	636.4	38.1	873	4 AAF81719	AAF81719 Human pro
31	636	38.1	925	6 ABL90148	ABL90148 Human pol
32	523.6	31.4	587	4 AAH12830	AAH12830 Human CDN
33	493.4	29.6	823	6 ABK30401	ABK30401 Human G-p
34	434.2	26.0	1083	4 AAC85697	AAC85697 Nucleotid
35	321	19.2	2751	6 AAD38311	AAD38311 Murine di
36	321	19.2	3287	6 AAD38955	AAD38955 Alternati
37	313	18.8	2495	6 AAD38957	AAD38957 Human dip
38	313	18.8	2617	6 ABK83323	ABK83323 cDNA enco
39	313	18.8	2660	6 ADI16689	ADI16689 Human NOV
40	313	18.8	2660	6 ADI16687	ADI16687 Human NOV
41	313	18.8	2660	12 ADN42341	ADN42341 Human CDN
42	313	18.8	3024	6 AAD38954	AAD38954 Human dip
43	313	18.8	3660	12 ADQ67504	ADQ67504 Novel hum
44	313	18.8	3716	6 ABQ75955	ABQ75955 Human PMW
45	313	18.8	4076	6 ABK83337	ABK83337 cDNA enco

ALIGNMENTS

RESULT 1  
AAC85696  
ID AAC85696 standard; cDNA; 1669 BP.  
XX  
AC AAC85696;  
XX  
DT 29-JUN-2001 (first entry)  
XX  
DE Nucleotide sequence encoding human DPP8 318Thr-658Val+759Ala-882Ile.  
XX  
KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;  
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;  
KW growth hormone deficiency; glucose level; mucosal regeneration;  
KW non-insulin dependent diabetes mellitus; glucose intolerance;  
KW immunosuppression; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 2..1399  
FT /tag= a  
FT /product= "Human DPP8 318Thr-658Val+759Ala-882Ile"

XX  
PN WO200119866-A1.  
XX  
PD 22-MAR-2001.  
XX  
PF 11-SEP-2000; 2000WO-AU001085.  
XX  
PR 10-SEP-1999; 99AU-00002762.  
PR 18-FEB-2000; 2000AU-00005709.  
XX  
PA (UNSY ) UNIV SYDNEY.  
XX  
PI Abbott CA, Gorell MD;  
XX  
DR WPI; 2001-281520/29.  
DR P-PSDB; AAB47189.  
XX  
PT New human dipeptidyl aminopeptidase (DPP8) useful for cleaving  
PT substrates, identifying inhibitors of DPP8 catalytic activity which have  
PT therapeutic uses, and for detecting activated T cells.

XX Claim 21; Page 72-74; 78pp; English.

XX The sequences given in AAC85695-97 encode fragments of human dipeptidyl

CC aminopeptidase (DPP8). DPP8 has substrate specificity for H-Ala-Pro-pNA,

CC H-Gly-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a prolyl oligopeptidase

CC and a dipeptidyl peptidase, because it is capable of hydrolysing the

CC peptide bond C-terminal to Pro in each of these compounds. DPP8 is

CC homologous with human DPP4. DPP8 is useful for cleaving a substrate, and

CC for detecting an activated T cell which involves measuring the level of

CC DPP8 gene expression in a T cell. The level of DPP8 expression is

CC detected by detecting the amount of DPP8 RNA in the cell. It is also

CC useful for identifying a molecule capable of inhibiting the cleavage of

CC the substrate by DPP8. Molecules identified as inhibiting DPP8 catalytic

CC activity may be useful for treating diarrhoea, growth hormone deficiency,

CC lowering glucose levels in non-insulin dependent diabetes mellitus and

CC other disorders involving glucose intolerance, enhancing mucosal

CC regeneration and as immunosuppressants

XX SQ Sequence 1669 BP; 524 A; 336 C; 336 G; 473 T; 0 U; 0 Other;

Query Match 100.0%; Score 1669; DB 4; Length 1669;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTCAGAAATATGATGCTGA 60

DB 1 AACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTCAGAAATATGATGCTGA 60

QY 61 AGNAGGATCATAGATGTCATAGATAGGAACTATTCACCTTTTGAGATCTCTTTTGA 120

DB 61 AGNAGGATCATAGATGTCATAGATAGGAACTATTCACCTTTTGAGATCTCTTTTGA 120

QY 121 AGGAGTTGAATATATTGCCAGAGCTGGATGGAATCTCCTGAGGAAATATGCTGGTCCAT 180

DB 121 AGGAGTTGAATATATTGCCAGAGCTGGATGGAATCTCCTGAGGAAATATGCTGGTCCAT 180

QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTAT 240

DB 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTAT 240

QY 241 CCCAGTAGAAGATGATGTTATGGAAGGAGGAGACTCAATGAGTCAGTGGCTGATTCGT 300

DB 241 CCCAGTAGAAGATGATGTTATGGAAGGAGGAGACTCAATGAGTCAGTGGCTGATTCGT 300

QY 301 GAGGCCACTAATTTATCTATGAAGAACACACAGACATCTGGATAAATATCCATGACATCTT 360

DB 301 GAGGCCACTAATTTATCTATGAAGAACACACAGACATCTGGATAAATATCCATGACATCTT 360

QY 361 TCATGTTTTTCCCAAGTCACGAGAGGAAATTCAGTATTATTTTGGCTCTGAATGCAA 420

DB 361 TCATGTTTTTCCCAAGTCACGAGAGGAAATTCAGTATTATTTTGGCTCTGAATGCAA 420

QY 421 AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATAACG 480

DB 421 AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATAACG 480

QY 481 ATCCAGTGGGGCTCGCTCCAGTGATTTCAAGTGCTCTATCAAGGAGGATAGC 540

DB 481 ATCCAGTGGGGCTCGCTCCAGTGATTTCAAGTGCTCTATCAAGGAGGATAGC 540

QY 541 AATTACAGTGGTGAATGGAGTTCTTGGCCGGATGATCTAATATCCAAAGTTGATGA 600

DB 541 AATTACAGTGGTGAATGGAGTTCTTGGCCGGATGATCTAATATCCAAAGTTGATGA 600

QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATACCTGTGA 660

DB 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATACCTGTGA 660

QY 661 CGTAGTCAGTTACGTAATCTGGAGAGGTGAACAAGGCTGACTGACCGTGCTACTCACA 720

DB 661 CGTAGTCAGTTACGTAATCTGGAGAGGTGAACAAGGCTGACTGACCGTGCTACTCACA 720

QY 721 TTCTTGCTGCATCAGTCAGCATGTGACCTTTTATTAAGTAAGTATAGTACACGAGAA 780

DB 721 TTCTTGCTGCATCAGTCAGCATGTGACCTTTTATTAAGTAAGTATAGTACACGAGAA 780

QY 781 TCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCTGAGAGTACGCAATCTCGAAGAC 840

DB 781 TCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCTGAGAGTACGCAATCTCGAAGAC 840

QY 841 AAAGGAAATTTGGGCGCACCATTTTGGATTCAGCAGGTCCTCTCTGACTATACCTCTCC 900

DB 841 AAAGGAAATTTGGGCGCACCATTTTGGATTCAGCAGGTCCTCTCTGACTATACCTCTCC 900

QY 901 AGAAATTTCTCTTTTGAAGTACTACTGGAATTCATGTATGGATGCTCTACAAGCC 960

DB 901 AGAAATTTCTCTTTTGAAGTACTACTGGAATTCATGTATGGATGCTCTACAAGCC 960

QY 961 TCATGATCTACAGCCTGGAAAGAAATATCCTACTCTGCTGTTCATATATGTTGCTCTCA 1020

DB 961 TCATGATCTACAGCCTGGAAAGAAATATCCTACTCTGCTGTTCATATATGTTGCTCTCA 1020

QY 1021 GGTGTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1080

DB 1021 GGTGTGCTATTTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1080

QY 1081 ACGTTATATGGGTCACTGAGAACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCAT 1140

DB 1081 ACGTTATATGGGTCACTGAGAACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCAT 1140

QY 1141 GCAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200

DB 1141 GCAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200

QY 1201 TGAGATGTCATTTTGGCAGTACATGATATTTACTGAGTTTTTTAGTGAGGGCTGGAAA 1260

DB 1201 TGAGATGTCATTTTGGCAGTACATGATATTTACTGAGTTTTTTAGTGAGGGCTGGAAA 1260

QY 1261 GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTGATTCGGG 1320

DB 1261 GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCTGATTCGGG 1320

QY 1321 AGAACATTTAGAACTGCACTCTTTTGGCACTACCTTCAAGAAAACCTTGGATCACGTAATGC 1380

DB 1321 AGAACATTTAGAACTGCACTCTTTTGGCACTACCTTCAAGAAAACCTTGGATCACGTAATGC 1380

QY 1381 TGCTCTAAAAAGTGATATAAATTTTGAACCTGTAGAACTCTCTGGTATACACCTGGCTATTT 1440

DB 1381 TGCTCTAAAAAGTGATATAAATTTTGAACCTGTAGAACTCTCTGGTATACACCTGGCTATTT 1440

QY 1441 AACCAAAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCATTTTGATACC 1500

DB 1441 AACCAAAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCATTTTGATACC 1500

QY 1501 TGCCATGTAACTACTCTCTGAAAAATAAATGTTGGTGCATGCGGGGTCTACGGTTTGT 1560

DB 1501 TGCCATGTAACTACTCTCTGAAAAATAAATGTTGGTGCATGCGGGGTCTACGGTTTGT 1560

QY 1561 GGTAGTAACTCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 1620

DB 1561 GGTAGTAACTCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 1620

QY 1621 GACCCAGCAATACCAATGAAGTACTTAAABAAAAAABAAAAAABAAAAA 1669

DB 1621 GACCCAGCAATACCAATGAAGTACTTAAABAAAAAABAAAAAABAAAAA 1669

RESULT 2

ABN59775

ID ABN59775 standard; cDNA; 2669 BP.

XX

AC ABN59775;

XX

DT 28-JUN-2002 (first entry)

XX





Db	2450	AACCAANTGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATCACATTITGATACC	2509	Db	1184	AACAGGTACAGCAATCCTTAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGA	1243
Qy	1501	TGCCATGTAACTACTCTCTGAAATAAATGTTGGTGCATGAGGGCTCTAGCTTTGT	1560	Qy	61	AGGAGGATCATAGATGTCATAGATAAGGAACTAAATTCOAACCTTTTGAGATCTTATTGA	120
Db	2510	TGCCATGTAACTACTCTCTGAAATAAATGTTGGTGCATGAGGGCTCTAGCTTTGT	2569	Db	1244	AGGAGGATCATAGATGTCATAGATAAGGAACTAAATTCOAACCTTTTGAGATCTTATTGA	1303
Qy	1561	GGTAGTAATCTAATACCTTTAAACCCACCATGCTCAAAATCAAAATGATATATTCCTTGAGA	1620	Qy	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAAATATGCTTGGTCCAT	180
Db	2570	GGTAGTAATCTAATACCTTTAAACCCACCATGCTCAAAATCAAAATGATATATTCCTTGAGA	2629	Db	1304	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAAATATGCTTGGTCCAT	1363
Qy	1621	GACCCAGCAATACCAATAGAAATTAATAAAAAA	1659	Qy	181	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATATTAT	240
Db	2630	GACCCAGCAATACCAATAGAAATTAATAAAAAA	2668	Db	1364	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATATTAT	1423
RESULT 3				Qy	241	CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTTGAGTCAGTCGCTGATTTCTGT	300
ABNS9774				Db	1424	CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTTGAGTCAGTCGCTGATTTCTGT	1483
AC	ABNS9774;			Qy	301	GAGGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCCATGACATCTT	360
XX	28-JUN-2002 (first entry)			Db	1484	GAGGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCCATGACATCTT	1543
DE	Novel human coding sequence SEQ ID NO: 185.			Qy	361	TCATGTTTTTCCCAAGTCACGAAGAGGAAATTCAGTTTATTTTGGCTCTGGAATGCAA	420
XX	Human; antianemic; vulnerary; antiinflammatory; immunomodulator;			Db	1544	TCATGTTTTTCCCAAGTCACGAAGAGGAAATTCAGTTTATTTTGGCTCTGGAATGCAA	1603
KW	antiinfertility; cerebroprotective; cytotestic; rheumatic; gene therapy;			Qy	421	AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAACAAATATAACG	480
KW	neuroprotective; antiparkinsonian; protein therapy; EST;			Db	1604	AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAACAAATATAACG	1663
XX	expressed sequence tag; gene; ss.			Qy	481	ATCCAGTGGTGGGCTGCTCCTCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC	540
OS	Homo sapiens.			Db	1664	ATCCAGTGGTGGGCTGCTCCTCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC	1723
PN	WO200222660-A2.			Qy	541	AATACAGTGGTGGATGGAAATTTCTTGGCCGATGGATCGATCTAATATCAAGTTGATGA	600
PD	21-MAR-2002.			Db	1724	AATACAGTGGTGGATGGAAATTTCTTGGCCGATGGATCTAATATCAAGTTGATGA	1783
XX	10-SEP-2001, 2001WO-US026015.			Qy	601	AGTCAGAGGCTGGTATATTTTGAAGGACCAAAGACTCCCTTTAGAGCATCACTGTGA	660
XX	11-SEP-2000, 2000US-00659671.			Db	1784	AGTCAGAGGCTGGTATATTTTGAAGGACCAAAGACTCCCTTTAGAGCATCACTGTGA	1843
PR	(HYSE-) HYSEQ INC.			Qy	661	CGTAGTCAGTTAGCTAAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA	720
XX	Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;			Db	1844	CGTAGTCAGTTAGCTAAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA	1903
PI	Xue AJ, Yang Y, Wehrman T, Drmanac RT;			Qy	721	TTCTTGTGCTGATCAGTCAGCTGCTGCTTTTAAAGTAAAGTATAGTAAACCAAGAGAA	780
XX	WPI; 2002-292408/33.			Db	1904	TTCTTGTGCTGATCAGTCAGCTGCTGCTTTTAAAGTAAAGTATAGTAAACCAAGAGAA	1963
XX	P-PSDB; ABNS97361.			Qy	781	TCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCACAAAC	840
PT	An isolated polynucleotide for treating diseases associated with its			Db	1964	TCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTGAAGATGACCCAACTTGCACAAAC	2023
PS	encoded polypeptide such as cancer and multiple sclerosis.			Qy	841	AAAGGAAATTTTGGGCCACCATTTTGGATTTGAGCAGGCTCTCTCTGACTATATCTCTCC	900
XX	Claim 1; SEQ ID NO 185; 509pp; English.			Db	2024	AAAGGAAATTTTGGGCCACCATTTTGGATTTGAGCAGGCTCTCTCTGACTATATCTCTCC	2083
CC	The present invention provides the protein and coding sequences of 444			Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAAGCC	960
CC	novel human proteins. These were isolated from expressed sequences tags			Db	2084	AGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTACAAGCC	2143
CC	(ESTs). They can be used to stimulate cell growth, to regulate			Qy	961	TCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTATATATGTTGCTCTCA	1020
CC	haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth			Db	2144	TCATGATCTACAGCTGGAAAGAAATATCTTACTGTGCTGTTTATATATGTTGCTCTCA	2203
CC	e.g. in burn treatment, to regulate the immune system e.g. to treat			Qy	1021	GGTTGCTATTGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATCATACAGATACACGGA	1080
CC	multiple sclerosis, to regulate activin or inhibin e.g. to treat			Db	2204	GGTTGCTATTGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATCATACAGATACACGGA	2263
CC	infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke			Qy	1081	ACGTTATATGGTCACTCTGACCAAGATGACAGGGCTATTACTAGGATCTGTGGCCAT	1140
CC	and cancer, to screen for drugs, to treat inflammatory conditions e.g.			Db	2264	ACGTTATATGGTCACTCTGACCAAGATGACAGGGCTATTACTTAGGATCTGTGGCCAT	2323
CC	rheumatoid arthritis, and to treat nervous system disorders e.g.						
CC	Parkinson's disease. The present sequence is a coding sequence of the						
XX	invention						
Qy	Sequence 2842 BP; 857 A; 592 C; 635 G; 758 T; 0 U; 0 Other;						
XX	Query Match 99.4%; Score 1659; DB 6; Length 2842;						
XX	Best Local Similarity 100.0%; Pred. No. 0;						
XX	Matches 1659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Qy	1 AACCGGTACAGCAATCCTTAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGA						
XX							

QY 1141 GCAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 1200  
DB 2324 GCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 2383  
QY 1201 TGAGAACTGTCATTTTGCACATACCATATATTACTAGTGTATTTTGTAGTGGGCTGGAAA 1260  
DB 2384 TGAGAACTGTCATTTTGCACATACCATATATTACTAGTGTATTTTGTAGTGGGCTGGAAA 2443  
QY 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTTCTGAATCGGG 1320  
DB 2444 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTTCTGAATCGGG 2503  
QY 1321 AGACATATGAACTGATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCACTGATTCG 1380  
DB 2504 AGACATATGAACTGATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCACTGATTCG 2563  
QY 1381 TGCTCTAAAGTGATATAATTTTGCCTGTGTAGAACTCTCTGTATACACTGGCTATTT 1440  
DB 2564 TGCTCTAAAGTGATATAATTTTGCCTGTGTAGAACTCTCTGTATACACTGGCTATTT 2623  
QY 1441 AACCAATAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCATTTTGTATACC 1500  
DB 2624 AACCAATAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCATTTTGTATACC 2683  
QY 1501 TGCCATGTAACATCTACTCTCTGAAAATAAATGTGGTGCATCCAGGGGTCTACGGTTTGT 1560  
DB 2684 TGCCATGTAACATCTACTCTCTGAAAATAAATGTGGTGCATCCAGGGGTCTACGGTTTGT 2743  
QY 1561 GGTAGTAACTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATTTCTTGAGA 1620  
DB 2744 GGTAGTAACTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATTTCTTGAGA 2803  
QY 1621 GACCAGCAATACCAATAGAAATTAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 1659  
DB 2804 GACCAGCAATACCAATAGAAATTAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 2842

RESULT 4

ID ABK83325 standard; cDNA; 4523 BP.  
XX ABK83325;

AC ABK83325;  
DT 12-AUG-2002 (first entry)  
DE cDNA encoding human DPP-1 splice variant #1.

KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPP;  
KW DPP-1; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;  
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;  
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;  
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;  
KW dyskinesia; reproductive disorder; inflammatory disorder;  
KW metabolic disorder; gene; ss.

XX Homo sapiens.  
OS  
XX

XX WO200231134-A2.  
FN

XX 18-APR-2002.  
PD

XX 12-OCT-2001; 2001WO-US031874.  
PF

XX 12-OCT-2000; 2000US-0240117P.  
PR

XX (FERR ) FERRING BV.  
PA

XX Qi S, Akinsanya KO, Riviere PJ, Junien J;  
PI

XX WPI; 2002-444178/47.  
DR

XX P-PSDB; ABG61594.  
DR

XX

PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding  
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and  
PT viral infections, cancers, allergies, neurological disorders, or pain.  
XX Disclosure; Page 61-62; 113pp; English.  
XX  
CC The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPP-IV)-related proteins  
CC (DPP-IV). The dipeptidyl peptidase IV-related proteins (DPP-IV) and nucleic  
CC acids encoding them are useful for treating infections such as fungal,  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurokinetic  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322  
CC -ABK83343 encode human DPP-IV proteins  
XX  
SQ Sequence 4523 BP; 1384 A; 828 C; 940 G; 1371 T; 0 U; 0 Other;  
Query Match 98.8%; Score 1648.4; DB 6; Length 4523;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;  
QY 1 AACAGGTACACCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATGCTGA 60  
DB 1164 AACAGGTACACCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATGCTGA 1223  
QY 61 AGGAAGGATCATAGATGTCATAGATAGGAATCAATCAACCTTTTGGATTTCTATTGA 120  
DB 1224 AGGAAGGATCATAGATGTCATAGATAGGAATCAATCAACCTTTTGGATTTCTATTGA 1283  
QY 121 AGGAGTTGAATATATTCGAGAGCTGGATGGACTCTGAGGGAATATGCTTGGTCCAT 180  
DB 1284 AGGAGTTGAATATATTCGAGAGCTGGATGGACTCTGAGGGAATATGCTTGGTCCAT 1343  
QY 181 CCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTGTGTGATCTCACCTGAATTTTAT 240  
DB 1344 CCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTGTGTGATCTCACCTGAATTTTAT 1403  
QY 241 CCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTTAGTGCAGTGCCTGATTTCTGT 300  
DB 1404 CCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTTAGTGCAGTGCCTGATTTCTGT 1463  
QY 301 GACGCCACTAATTTATCTATGAAGAACACAGACATCTCGATTAATATCCATGACATCTT 360  
DB 1464 GACGCCACTAATTTATCTATGAAGAACACAGACATCTCGATTAATATCCATGACATCTT 1523  
QY 361 TCATGTTTTTCCCAAGTCCAGAGGAAATTTGAGTTTATTTTGGCTCTGATGCAA 420  
DB 1524 TCATGTTTTTCCCAAGTCCAGAGGAAATTTGAGTTTATTTTGGCTCTGATGCAA 1583  
QY 421 AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAAG 480  
DB 1584 AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAAG 1643  
QY 481 ATCCAGTGTGGGCTGCTGCTCAAGTGTATTTCAAGTGTCTATCAAGAGAGATAGC 540  
DB 1644 ATCCAGTGTGGGCTGCTGCTCAAGTGTATTTCAAGTGTCTATCAAGAGAGATAGC 1703  
QY 541 AATTACCAGTGTGATGGAAGTCTTCCGCGGATGATCTTAATATCCAGTTGATGA 600  
DB 1704 AATTACCAGTGTGATGGAAGTCTTCCGCGGATGATCTTAATATCCAGTTGATGA 1763  
QY 601 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCAGCTGA 660  
DB 1764 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCAGCTGA 1823  
QY 661 CGTAGTCAGTTACGTAATCTCGAGAGGTGAACAGGCTGACTACCGTGGCTACTCACA 720  
|||||



QY	181	CCTACTAGATGCTCCAGACTCGCCTACAGATAGTGTGATCTCACTGAATTAATTTAT	240
Db	1183	CCTACTAGATGCTCCAGACTCGCCTACAGATAGTGTGATCTCACTGAATTAATTTAT	1242
QY	241	CCCAAGTAAAGATGATGTTATGAAAGGAGGAGAGACTCAATGAGTCAGTGCCTGATTCGT	300
Db	1243	CCCAAGTAAAGATGATGTTATGAAAGGAGGAGAGACTCAATGAGTCAGTGCCTGATTCGT	1302
QY	301	GACGCCACTAATTTATCTATGAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	360
Db	1303	GACGCCACTAATTTATCTATGAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	1362
QY	361	TCATGTTTTTCCCAAGAGTCACGAAGGAAATGAGTTATTTTGGCTCTGAATGCAA	420
Db	1363	TCATGTTTTTCCCAAGAGTCACGAAGGAAATGAGTTATTTTGGCTCTGAATGCAA	1422
QY	421	AACAGGTTTCGCTCATTTATACAAATTAATCTATTTTAAAGGAAACAAATATAACG	480
Db	1423	AACAGGTTTCGCTCATTTATACAAATTAATCTATTTTAAAGGAAACAAATATAACG	1482
QY	481	ATCCAGTGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGC	540
Db	1483	ATCCAGTGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGC	1542
QY	541	AATTACAGTGTGGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATCCAAGTTGATGA	600
Db	1543	AATTACAGTGTGGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATCCAAGTTGATGA	1602
QY	601	AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTA	660
Db	1603	AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTA	1662
QY	661	CGTAGTCAGTTACGTAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGCTACTCACA	720
Db	1663	CGTAGTCAGTTACGTAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGCTACTCACA	1722
QY	721	TTCTTGCTGTCAGTCAGTCAGCTGCTCTTTTAAAGTAAAGTAAAGTAAAGTAAAGTAA	780
Db	1723	TTCTTGCTGTCAGTCAGTCAGCTGCTCTTTTAAAGTAAAGTAAAGTAAAGTAAAGTAA	1782
QY	781	TCCACACTGTGTCCTTTACAGCTATCAAGTCTCAAGTCTCAAGTCAAGTCTCAAGTCA	840
Db	1783	TCCACACTGTGTCCTTTACAGCTATCAAGTCTCAAGTCTCAAGTCAAGTCTCAAGTCA	1839
QY	841	AAAGGAATTTTGGGCCACCAATTTTGATTTGAGTCTCTCTCTGACTATATCTCTCC	900
Db	1840	AAAGGAATTTTGGGCCACCAATTTTGATTTGAGTCTCTCTCTGACTATATCTCTCC	1899
QY	901	AGAAATTTTCTTTTGAAGTACTACTGGAATTAATGATGGAATGCTCTTCAAGCC	960
Db	1900	AGAAATTTTCTTTTGAAGTACTACTGGAATTAATGATGGAATGCTCTTCAAGCC	1959
QY	961	TCATGATCTACAGCTGGAAGAAATATCCTACTGCTGCTGCTATATATGCTGCTCTCA	1020
Db	1960	TCATGATCTACAGCTGGAAGAAATATCCTACTGCTGCTGCTATATATGCTGCTCTCA	2019
QY	1021	GTTGCTATTTGCTGGGCCCCAGTCACTCTGTTGATCTTCTATGATACAGATACACGGA	1080
Db	2020	GTTGCTATTTGCTGGGCCCCAGTCACTCTGTTGATCTTCTATGATACAGATACACGGA	2079
QY	1081	ACGTTATATGGGTCACTGACCAAGTAAACAGGGCTATTACTTATGATCTGTGGCCAT	1140
Db	2080	ACGTTATATGGGTCACTGACCAAGTAAACAGGGCTATTACTTATGATCTGTGGCCAT	2139
QY	1141	GCAAGCAGAAAGTTCCTCTGACCAAAATCGTTTACTGCTTCTTACATGTTTCTTGGGA	1200
Db	2140	GCAAGCAGAAAGTTCCTCTGACCAAAATCGTTTACTGCTTCTTACATGTTTCTTGGGA	2199
QY	1201	TGAGAATGTCATTTTGGCATACACCAAGTATATTTAGTGGTCTGAGGCTGGAAA	1260
Db	2200	TGAGAATGTCATTTTGGCATACACCAAGTATATTTAGTGGTCTGAGGCTGGAAA	2259
QY	1261	GCCATATGATTTACAGATCTATCTCCTCAGGAGAGACACAGCATTAAGATTCTCTGAATCGGG	1320

Db	2260	GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAAGATTCTCTGAATCGGG	2319
QY	1321	AGAACATTTATGAACCTGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTTATGC	1380
Db	2320	AGAACATTTATGAACCTGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTTATGC	2379
QY	1381	TGCTCTAAAAAGTGATATAAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	1440
Db	2380	TGCTCTAAAAAGTGATATAAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	2439
QY	1441	BAACAAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCAATTTTATACAC	1500
Db	2440	BAACAAATGAGGAGGTTTAAATCAACAGAAAAACAGAAATTTGATCATCAATTTTATACAC	2499
QY	1501	TGCCATGTAAACATCTACTCTCTGAAATAAATGTGGTGCATGCAGGGGTCTACGGTTTGT	1560
Db	2500	TGCCATGTAAACATCTACTCTCTGAAATAAATGTGGTGCATGCAGGGGTCTACGGTTTGT	2559
QY	1561	GGTAGTAAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTCCTGAGA	1620
Db	2560	GGTAGTAAATCTAATACCTTAAACCCCATGCTCAAAATCAAAATGATACATATTCCTGAGA	2619
QY	1621	GACCCAGCAATACCATTAAGAAATTTACTAA	1648
Db	2620	GACCCAGCAATACCATTAAGAAATTTACTAA	2647
RESULT 6			
ABK83331			
ID	ABK83331	standard; cDNA; 4676 BP.	
XX	AC	ABK83331;	
XX	AC	ABK83331;	
DT	12-AUG-2002	(first entry)	
DE	cDNA encoding human DPRP-1 splice variant #7.		
XX	Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;		
XX	DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;		
XX	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;		
XX	heart failure; hypertension; urinary retention; osteoporosis; cancer;		
XX	ulcer; allergy; cancer; psychotic disorder; neurological disorder;		
XX	dyskinesia; reproductive disorder; inflammatory disorder;		
XX	metabolic disorder; gene; ss.		
OS	Homo sapiens.		
XX	WO200231134-A2.		
XX	18-APR-2002.		
PF	12-OCT-2001; 2001WO-US031874.		
XX	12-OCT-2000; 2000US-0240117P.		
PR	(FERR ) FERRING BV.		
PA	Qi S, Akinsanya KO, Riviere PJ, Junien J;		
XX	WPI; 2002-444178/47.		
DR	P-PSDB; ABG61600.		
XX	New dipeptidyl peptidase IV-related proteins and nucleic acids encoding		
PT	the proteins, useful for treating e.g. fungal, bacterial, protozoan and		
PT	viral infections, cancers, allergies, neurological disorders, or pain.		
XX	Disclosure; Page 72-73; 113pp; English.		
XX	The present invention relates to the isolation of novel human serine		
CC	proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins		
CC	(DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic		
CC	acids encoding them are useful for treating infections such as fungal,		

CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinetic  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322  
CC -ABK83343 encode human DPRP proteins  
XX  
SQ Sequence 4676 BP; 1424 A; 859 C; 979 G; 1414 T; 0 U; 0 Other;

Query Match 89.9%; Score 1501; DB 6; Length 4676;  
Best Local Similarity 91.5%; Pred. No. 0;  
Matches 1663; Conservative 5; Mismatches 149; Gaps 1;  
  
QY 1 AACAGGTACAGCAATCTCTAAGTCTACCTTTTAAAGTCTCAGAAATATGATCTCATCTGA 60  
DB 1164 AACAGGTACAGCAATCTCTAAGTCTACCTTTTAAAGTCTCAGAAATATGATCTCATCTGA 1223  
  
QY 61 AGAAGGATCATAGATGTCATAGATAGCAATTAAGCACTTTTGAAGTCTCAGAAATATGATCTCATCTGA 120  
DB 1224 AGAAGGATCATAGATGTCATAGATAGCAATTAAGCACTTTTGAAGTCTCAGAAATATGATCTCATCTGA 1283  
  
QY 121 AGAGTGTGAATATATGCGAGAGCTGGATGCACTCTGAGGGAATATGCTGGTCCAT 180  
DB 1284 AGAGTGTGAATATATGCGAGAGCTGGATGCACTCTGAGGGAATATGCTGGTCCAT 1343  
  
QY 181 CCTACATAGCTGCTCCAGACTCGGCTACAGATAGTGTGATCTCAGCTGAATATTATAT 240  
DB 1344 CCTACATAGCTGCTCCAGACTCGGCTACAGATAGTGTGATCTCAGCTGAATATTATAT 1403  
  
QY 241 CCCAGTAGAAGATGATGTTATGAAAGGACAGAGACTCAATGAGTCAGTGCCTGATCTGT 300  
DB 1404 CCCAGTAGAAGATGATGTTATGAAAGGACAGAGACTCAATGAGTCAGTGCCTGATCTGT 1463  
  
QY 301 GAGCCACTAATATCTATGAAGAAACACAGACATCTGGAATATATCCATGACATCTT 360  
DB 1464 GAGCCACTAATATCTATGAAGAAACACAGACATCTGGAATATATCCATGACATCTT 1523  
  
QY 361 TCATGTTTTTCCCAAGTACAGAGAGAAATGAGTTTATTTTTGCCCTCTGAATGCAA 420  
DB 1524 TCATGTTTTTCCCAAGTACAGAGAGAAATGAGTTTATTTTTGCCCTCTGAATGCAA 1583  
  
QY 421 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAAG 480  
DB 1584 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATAAAG 1643  
  
QY 481 ATCCAGTGTGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGC 540  
DB 1644 ATCCAGTGTGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGC 1703  
  
QY 541 AATTACAGTGTGATGGAGATCTTGGCGGCGATGGATCTAATATCCAAGTTCATGA 600  
DB 1704 AATTACAGTGTGATGGAGATCTTGGCGGCGATGGATCTAATATCCAAGTTCATGA 1763  
  
QY 601 AGTCAGAAGCTGGTATATTTTGAAGGCAACAAAGATCCCTCTTTAGAGCATCACCTGTA 660  
DB 1764 AGTCAGAAGCTGGTATATTTTGAAGGCAACAAAGATCCCTCTTTAGAGCATCACCTGTA 1823  
  
QY 661 CGTAGTCAGTTACGTAATATCTGGAGAGGTGACAGGCTGATCCGCTGGCTACTACA 720  
DB 1824 CGTAGTCAGTTACGTAATATCTGGAGAGGTGACAGGCTGATCCGCTGGCTACTACA 1883  
  
QY 721 TTCTTCTGATCAGTCAGTACACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAGAA 780  
DB 1884 TTCTTCTGATCAGTCAGTACACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAGAA 1943  
  
QY 781 TCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAAAAC 840  
DB 1944 TCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAAAAC 2003

QY 841 AAAAGAAATTTTGGGCAACCAATTTTGGATTTAGCAGGTCCTCTTCTGACTATATCTCTCC 900  
DB 2004 AAAAGAAATTTTGGGCAACCAATTTTGGATTTAGCAGGTCCTCTTCTGACTATATCTCTCC 2063  
  
QY 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCC 960  
DB 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCC 2123  
  
QY 961 TCATGATCTACAGCCCTGGAAAGAAATATCTTACTGTCTGTTCATATATGGTGGTC --- 1016  
DB 2124 TCATGATCTACAGCCCTGGAAAGAAATATCTTACTGTCTGTTCATATATGGTGGTCGGGT 2183  
  
QY 1017 ----- 1016  
DB 2184 CAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTC 2243  
  
QY 1017 ----- 1016  
DB 2244 ATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATG 2303  
  
QY 1017 -----CTCAGGTTGCTATTGCTGGGGCCCGCAGTCACCTCTG 1051  
DB 2304 GCATTAATGCAGGTCAGATATCTTCAGGGTTCCTATTGCTGGGGCCCGCAGTCACCTCTG 2363  
  
QY 1052 TCGATCTTCTATGATACAGGATACACGGACGTTTATATGGGTACCCCTGACCCAGAAATGAA 1111  
DB 2364 TGGATCTTCTATGATACAGGATACACGGACGTTTATATGGGTACCCCTGACCCAGAAATGAA 2423  
  
QY 1112 CAGGGCTATTAATTAGGATCTGTGGCATTCAGGAGAAAGTTCCCTCTGACCAAT 1171  
DB 2424 CAGGGCTATTAATTAGGATCTGTGGCATTCAGGAGAAAGTTCCCTCTGACCAAT 2483  
  
QY 1172 CGTTTACTGCTCTTACATGCTTCTGATGAGAAATGTCATTTTGCACATACCAATATA 1231  
DB 2484 CGTTTACTGCTCTTACATGCTTCTGATGAGAAATGTCATTTTGCACATACCAATATA 2543  
  
QY 1232 TTACTGAGTTTTTATGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCAGAG 1291  
DB 2544 TTACTGAGTTTTTATGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCAGAG 2603  
  
QY 1292 AGACACAGCATAGAGTTCCTGAATCGGAGAACATTTAAGAACTGCAATTTTGCACATC 1351  
DB 2604 AGACACAGCATAGAGTTCCTGAATCGGAGAACATTTAAGAACTGCAATTTTGCACATC 2663  
  
QY 1352 CTTCAAGAAACCTTGGATCAGTATTTGCTCTTAAAGTGATATAATTTTGACCTGTG 1411  
DB 2664 CTTCAAGAAACCTTGGATCAGTATTTGCTCTTAAAGTGATATAATTTTGACCTGTG 2723  
  
QY 1412 TAGAACTCTCTGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAA 1471  
DB 2724 TAGAACTCTCTGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAAA 2783  
  
QY 1472 CACAGAAATGATCATCATCTTTGATACCTGCCATGTAACATCTACTCTGAAATATAAT 1531  
DB 2784 CACAGAAATGATCATCATCTTTGATACCTGCCATGTAACATCTACTCTGAAATATAAT 2843  
  
QY 1532 GTGGTCCATGAGGGTCTAGGTTTGGTGTAGTAATCTAATACCTTAACCCACATGC 1591  
DB 2844 GTGGTCCATGAGGGTCTAGGTTTGGTGTAGTAATCTAATACCTTAACCCACATGC 2903  
  
QY 1592 TCAAAATCAATATGATCATATTTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAA 1651  
DB 2904 TCAAAATCAATATGATCATATTTCTGAGAGACCCAGCAATACCATAGAAATTTACTAAAA 2963  
  
QY 1652 AAAAAAAAAAAAAA 1668  
DB 2964 AAAAAAAAAAAAAAGA 2980  
  
RESULT 7  
ACA92421  
ID ACA92421 standard; DNA; 2952 BP.  
XX

AC ACA92421;  
XX 15-JUL-2003 (first entry)  
XX  
DE DNA encoding human PMMM-6.  
XX  
XX Human; protein modification and maintenance molecule; PMMM; cancer;  
KW cell proliferation disorder; atherosclerosis; neurological disorder;  
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;  
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;  
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;  
KW infection; cycostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;  
KW antiinflammatory; thyromimetic; gene; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO2003031939-A2.  
XX  
XX PD 17-APR-2003.  
XX  
XX PF 11-OCT-2002; 2002WO-US032850.  
XX  
XX PR 12-OCT-2001; 2001US-0329689P.  
XX PR 25-OCT-2001; 2001US-0335703P.  
XX PR 09-NOV-2001; 2001US-0348887P.  
XX PR 28-NOV-2001; 2001US-0334145P.  
XX PR 06-DEC-2001; 2001US-0337451P.  
XX PR 14-DEC-2001; 2001US-0340584P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
PI Tran UK, B  cha SD, Duggan BM, Lee EA, Griffin JA, Li JX;  
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;  
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;  
PI Bhatia U, Burrill JJ, Lee S, Blake JJ, Ho A, Zheng W;  
XX  
XX WPI; 2003-430274/40.  
DR P-PSDB; ABU92026.  
XX  
XX PT New human protein modification and maintenance molecules (PMM) , useful  
PT for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or  
PT infections.  
XX  
XX PS Claim 5; Page 285-286; 31pp; English.  
XX  
XX CC The present invention relates to the isolation of human protein  
CC modification and maintenance molecules (PMM) , and the polynucleotide  
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM  
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful  
CC for diagnosing a condition or disease associated with the expression of  
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and  
CC generating an expression profile of a sample containing the  
CC polynucleotides. The diseases or conditions associated with decreased  
CC expression or overexpression of PMMM are cell proliferation disorders  
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,  
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
CC syndrome), gastrointestinal or epithelial disorders, and infections. The  
CC PMMM polypeptides or their fragments are useful in screening compounds  
CC for effectiveness as agonists or antagonists of the polypeptides, or in  
CC altering the expression of the target polynucleotide and compounds that  
CC specifically bind to, or modulate the activity of the polypeptide.  
XX ACA92416-ACA92455 encode the human PMMM polypeptides of the invention  
XX  
SQ Sequence 2952 BP; 886 A; 604 C; 664 G; 798 T; 0 U; 0 Other;

Query Match 89.5%; Score 1493.4; DB 10; Length 2952;  
Best Local Similarity 91.8%; Pred. No. 0;  
Matches 1651; Conservative 0; Mismatches 1; Indels 147; Gaps 1;

Qy 1 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATGTGAGAAATATGATTGATGCTGA 60  
Db |||||  
1154 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATGTGAGAAATATGATTGATGCTGA 1213  
Qy |||||  
61 AGAAGGATCATAGATGTCATAGATAAGAACTAATCAACCTTTTCAGATTCCTATTGGA 120  
Db |||||  
1214 AGAAGGATCATAGATGTCATAGATAAGAACTAATCAACCTTTTCAGATTCCTATTGGA 1273  
Qy |||||  
121 AGAGTTGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAAAAATATGCTTGGTCCAT 180  
Db |||||  
1274 AGAGTTGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAAAAATATGCTTGGTCCAT 1333  
Qy |||||  
181 CCTACTAGATCGCTCCAGATCGCCTACAGATAGTGTGATCTCAGCTGAATATTATTAT 240  
Db |||||  
1334 CCTACTAGATCGCTCCAGATCGCCTACAGATAGTGTGATCTCAGCTGAATATTATTAT 1393  
Qy |||||  
241 CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATTCAGTCAGTGCCTGATTCGT 300  
Db |||||  
1394 CCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCAATTCAGTCAGTGCCTGATTCGT 1453  
Qy |||||  
301 GAGCCCACTAATTTCTATGAAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360  
Db |||||  
1454 GAGCCCACTAATTTCTATGAAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1513  
Qy |||||  
361 TCATGTTTTTCCCAAGTCACGAAGAGAAATTGAGTTATTTTGGCTCTGATGCAA 420  
Db |||||  
1514 TCATGTTTTTCCCAAGTCACGAAGAGAAATTGAGTTATTTTGGCTCTGATGCAA 1573  
Qy |||||  
421 AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG 480  
Db |||||  
1574 AACAGGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACG 1633  
Qy |||||  
481 ATCCAGTGTGGGCTGCCTGCCCTCCAAAGTGAATTTCAAAGTGTCTATCAAGAGGAGATAGC 540  
Db |||||  
1634 ATCCAGTGTGGGCTGCCTGCCCTCCAAAGTGAATTTCAAAGTGTCTATCAAGAGGAGATAGC 1693  
Qy |||||  
541 AATTACAGTGTGTAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCCAAAGTTGATGA 600  
Db |||||  
1694 AATTACAGTGTGTAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCCAAAGTTGATGA 1753  
Qy |||||  
601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATACCTGTA 660  
Db |||||  
1754 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATACCTGTA 1813  
Qy |||||  
661 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACA 720  
Db |||||  
1814 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACA 1873  
Qy |||||  
721 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACAGAGAA 780  
Db |||||  
1874 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACAGAGAA 1933  
Qy |||||  
781 TCACACTGTGTGCTCCCTTTAAGCTATCAAGTCTCAAGATGACCCCACTGCGAAAC 840  
Db |||||  
1934 TCACACTGTGTGCTCCCTTTAAGCTATCAAGTCTCAAGATGACCCCACTGCGAAAC 1993  
Qy |||||  
841 AAAGGAATTTTGGGCGCACCATTTTGATTCAGCAGGCTCTCTTCTGACTATATCTCTCC 900  
Db |||||  
1994 AAAGGAATTTTGGGCGCACCATTTTGATTCAGCAGGCTCTCTTCTGACTATATCTCTCC 2053  
Qy |||||  
901 AGAAATTTTCTCTTTTGAAGAGTACTACTGAAATTTACATTTGATGGATGCTCTCAAGCC 960  
Db |||||  
2054 AGAAATTTTCTCTTTTGAAGAGTACTACTGAAATTTACATTTGATGGATGCTCTCAAGCC 2113  
Qy |||||  
961 TCATGATCTACAGCTCGAAAGAAATATCTTACTGCTGTGTTTCATATATGCTGCTCTCA 1020  
Db |||||  
2114 TCATGATCTACAGCTCGAAAGAAATATCTTACTGCTGTGTTTCATATATGCTGCTCTCA 2173  
Qy |||||  
1021 ----- 1020  
Db GGTGAGTGTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGC 2233  
Qy ----- 1020



Db 2234 CTCTCTAGGTTATGTGTTGTAGTATGATAGACAACAGGGGATCTGTGTCACCGAGGCTTAA 2293  
Qy 1021 -----GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTG 1053  
Db 2294 ATTTGAAGGGCCTTTAAATATATAATGTTGCTATTGCTGGGGCCCCAGTCACTCTGTG 2353  
Qy 1054 GATCTTCTATGTATACAGGATACACGGAACGTTTATATGTTGTCACCTGACCCAGAAATGAACA 1113  
Db 2354 GATCTTCTATGTATACAGGATACACGGAACGTTTATGTTGTCACCTGACCCAGAAATGAACA 2413  
Qy 1114 GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGCAACCAATCG 1173  
Db 2414 GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGCAACCAATCG 2473  
Qy 1174 TTATCTGCTCTTACATGGTTTCTGGATGAGAATGTCATTTTGCATACATACCAATATATT 1233  
Db 2474 TTATCTGCTCTTACATGGTTTCTGGATGAGAATGTCATTTTGCATACATACCAATATATT 2533  
Qy 1234 ACTGATCTTTTATGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAG 1293  
Db 2534 ACTGATCTTTTATGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAG 2593  
Qy 1294 ACACAGCATAAAGATGTTCTGTAATCGGAGAACATTATGAATGCTATCTTTTGCACCTACCT 1353  
Db 2594 ACACAGCATAAAGATGTTCTGTAATCGGAGAACATTATGAATGCTATCTTTTGCACCTACCT 2653  
Qy 1354 TCAAGAAAACCTTTGGATCAGCTATTGCTGCTCTAAAGTGCATATAATTTGACCTGTGA 1413  
Db 2654 TCAAGAAAACCTTTGGATCAGCTATTGCTGCTCTAAAGTGCATATAATTTGACCTGTGA 2713  
Qy 1414 GAATCTCTGTTATACATGCTGCTTTTAAACCAATGAGGAGTTTAACTCAACAGAAACA 1473  
Db 2714 GAATCTCTGTTATACATGCTGCTTTTAAACCAATGAGGAGTTTAACTCAACAGAAACA 2773  
Qy 1474 GAGATTTGATCATCATCTTTGATACCTGCGCATGTAACATCTACTCTCTGAAATTAATGT 1533  
Db 2774 GAGATTTGATCATCATCTTTGATACCTGCGCATGTAACATCTACTCTCTGAAATTAATGT 2833  
Qy 1534 GTGGCATCGAGGGCTTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTC 1593  
Db 2834 GTGGCATCGAGGGCTTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTC 2893  
Qy 1594 AAAATCAAATGATACATATCTCTGAGAGCCGACCAATACCAATGAATTACTAAAAA 1652  
Db 2894 AAAATCAAATGATACATATCTCTGAGAGCCGACCAATACCAATGAATTACTAAAAA 2952

RESULT 8  
AAH15009  
ID AAH15009 standard; cDNA; 2161 BP.  
XX

AC AAH15009;

DT 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:12963.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

KW Homo sapiens.

OS Homo sapiens.

PN EP1074617-A2.

XX 07-FEB-2001.

PF 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 12963; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides; and the combination of the 5'-end sequence and the 3'-end sequence is selected from those defined in the specification. The primer sets are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any special methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 2161 BP; 662 A; 429 C; 452 G; 618 T; 0 U; 0 Other;

Qy Query Match 89.1%; Score 1487.4; DB 4; Length 2161;

Db Best Local Similarity 91.7%; Pred. No. 0;

Qy Matches 1645; Conservative 0; Mismatches 1; Indels 147; Gaps 1;

Db 1 AACAGGTACAGCAAAATCTAAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGCTGA 60

369 AACAGGTACAGCAAAATCTAAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGCTGA 428

61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTTCAACCTTTTGAGATTTCTATTGA 120

429 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTTCAACCTTTTGAGATTTCTATTGA 488

121 AGGAGTTGAATATATTGGCCAGAGCTGGATGACCTCTCTGAGGAAATATGCTTGGTCCAT 180

489 AGGAGTTGAATATATTGGCCAGAGCTGGATGACCTCTCTGAGGAAATATGCTTGGTCCAT 548

181 CCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTGTGATCTCAGCTGAATATTAT 240

549 CCTACTAGATCGCTCCAGAGCTCGCTACAGATAGTGTGATCTCAGCTGAATATTAT 608

241 CCCAGTAGAAGATGATGTTATGGAAAGCAGAGACTCAATTTGAGTGCAGTGCATCTGT 300

609 CCCAGTAGAAGATGATGTTATGGAAAGCAGAGACTCAATTTGAGTGCAGTGCATCTGT 668

301 GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360

669 GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 728

361 TCATGTTTTTCCCAAGTCAAGAGGAAATTTGAGTTATTTTTTGCCTCTGAAATGCAA 420

729 TCATGTTTTTCCCAAGTCAAGAGGAAATTTGAGTTATTTTTTGCCTCTGAAATGCAA 788

421 AACAGGTTTCGCTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATATAACG 480



Db 789 AACAGGTTTCCTGCTATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAAGC 848  
Qy 481 ATCCAGTGTGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGC 540  
Db 849 ATCCAGTGTGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTCAAGAGGAGATAGC 908  
Qy 541 AATTACAGTGTGATGGAGTTCTTGGCCGCGCATGGATCTAAATATCAAGTTGATCA 600  
Db 909 AATTACAGTGTGATGGAGTTCTTGGCCGCGCATGGATCTAAATATCAAGTTGATCA 968  
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTTAGAGCATCACCTGTA 660  
Db 969 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTTAGAGCATCACCTGTA 1028  
Qy 661 CGTAGTCACTTACGTAATCTTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCA 720  
Db 1029 CGTAGTCACTTACGTAATCTTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCA 1088  
Qy 721 TTCTTGCTGCATCAGTCAGCTGACTGCTCTTTTAAAGTAAGTATAGTATACCAAGAA 780  
Db 1089 TTCTTGCTGCATCAGTCAGCTGACTGCTCTTTTAAAGTAAGTATAGTATACCAAGAA 1148  
Qy 781 TCCACACTGTGTCCCTTTCAAGCTATCAAGTCTTGAAGATGACCCAACTTGCRAAC 840  
Db 1149 TCCACACTGTGTCCCTTTCAAGCTATCAAGTCTTGAAGATGACCCAACTTGCRAAC 1208  
Qy 841 AAGGAATTTTGGCCACCATTTTGGATTTCAGAGGTCCTCTCTGACTATACTCTCC 900  
Db 1209 AAGGAATTTTGGCCACCATTTTGGATTTCAGAGGTCCTCTCTGACTATACTCTCC 1268  
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTTGAATTTTACATTTGTATGGATGCTCTCAAGCC 960  
Db 1269 AGAAATTTTCTCTTTTGAAGTACTACTTGAATTTTACATTTGTATGGATGCTCTCAAGCC 1328  
Qy 961 TCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTCATATATGTTGGTCTCA 1020  
Db 1329 TCATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTCATATATGTTGGTCTCA 1388  
Qy 1021 ----- 1020  
Db 1389 GGTGACAGTGTGTAATTCGGTTTAAAGAGTCAAGTATTTTCGGTTGAATACCTAGC 1448  
Qy 1021 ----- 1020  
Db 1449 CTCTCTAGTTATGTGTTGTAGTATAGACAAACAGGGGATCTGTACCGAGGGTTAA 1508  
Qy 1021 -----GTTGCTATTGCTGGGCCCCAGTCACTCTGTG 1053  
Db 1509 ATTGAAGGCGCTTTAATATAAATAGTTGCTATTGCTGGGCCCCAGTCACTCTGTG 1568  
Qy 1054 GATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACCAAGAAATGA 1113  
Db 1569 GATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACCAAGAAATGA 1628  
Qy 1114 GGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCG 1173  
Db 1629 GGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCG 1688  
Qy 1174 TTCTAGCTCTTACATGGTTTCTGGATGAGATGTCATTTTGGCATACCAATGATATT 1233  
Db 1689 TTCTAGCTCTTACATGGTTTCTGGATGAGATGTCATTTTGGCATACCAATGATATT 1748  
Qy 1234 ACTGAGTTTATAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAG 1293  
Db 1749 ACTGAGTTTATAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAG 1808  
Qy 1294 ACACAGCATAGAGTTCTCGAATCGGAGAACATTATGAATGCTATCTTTTGGCACTACCT 1353  
Db 1809 ACACAGCATAGAGTTCTCGAATCGGAGAACATTATGAATGCTATCTTTTGGCACTACCT 1868  
Qy 1354 TCAAGAAACCTTTGGATACGATTTGCTGCTCTAAAGTGATATAAATTTTGACCTGTGTA 1413

Db 1869 TCAAGAAACCTTTGGATCAAGTATTGCTGCTCTAAAGTGATATAAATTTTGACCTGTGTA 1928  
Qy 1414 GAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAACA 1473  
Db 1929 GAACTCTCTGGTATACACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAACA 1988  
Qy 1474 CAGAATTGATCATCATATTTTGATACCTGCGCATCTCTCTGAAATPAAATGT 1533  
Db 1989 CAGAATTGATCATCATATTTTGATACCTGCGCATCTCTCTGAAATPAAATGT 2048  
Qy 1534 GGTGCCATGAGGGGCTACGGTTTGTGGTAGTAAATCTAAATACCTTAACCCCATGCTC 1593  
Db 2049 GGTGCCATGAGGGGCTACGGTTTGTGGTAGTAAATCTAAATACCTTAACCCCATGCTC 2108  
Qy 1594 AAAATCAATGATACATATTCCTGAGAGCCAGCCAGCATACCATAGAAATTA 1646  
Db 2109 AAAATCAATGATACATATTCCTGAGAGCCAGCCAGCATACCATAGAAATTA 2161

RESULT 9  
ADV43982  
ID ADV43982 standard; cDNA; 2349 BP.  
XX  
AC ADV43982;  
XX  
DT 10-MAR-2005 (first entry)  
XX  
Human psychoneuroendocrine immune expressed sequence tag SEQ ID NO 1610.  
XX  
microarray; psychoneuroendocrine immune; chronic fatigue;  
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;  
KW cancer; neoplasm; infection; expressed sequence tag; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO2004108899-A2.  
XX  
PD 16-DEC-2004.  
XX  
PF 04-JUN-2004; 2004WO-US017686.  
XX  
PR 04-JUN-2003; 2003US-0475915P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Nicholson A, Vernon SD;  
XX  
DR WPI; 2005-031682/03.  
XX  
PT New microarray comprising probes for genes involved in  
PT psychoneuroendocrine immune (PNI) activity, useful in diagnosing a  
PT condition associated with PNI activity, e.g., inflammatory or infectious  
PT diseases.  
XX  
PS Claim 1; SEQ ID NO 1610; 254pp; English.  
XX  
The invention relates to a new microarray which comprises probes for  
CC genes involved in psychoneuroendocrine immune (PNI) activity. The  
CC microarray is useful in diagnosing a condition associated with PNI  
CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,  
CC cancer and infection. The present sequence represents a  
CC psychoneuroendocrine immune gene expressed sequence tag. Note the  
CC specificatio mentions SEQ ID NO of up to 3314 but only sequences up to  
CC SEQ ID NO 1829 are provided.  
XX  
SQ Sequence 2349 BP; 723 A; 459 C; 508 G; 659 T; 0 U; 0 Other;  
Query Match : 83.8%; Score 1399; DB 14; Length 2349;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAAGTGTACAGAAATATGATGCTGTA 60  
|||||

951 AACAGGTACAGCAAACTCTAAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGCTGA 1010  
61 AGAAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTTCAGATTCTATTGGA 120  
1011 AGAAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTTCAGATTCTATTGGA 1070  
121 AGAGTTGATATATTGCGCAGAGCTGGATGAGCTCCTGAGGGAATAATGCTGGTCCAT 180  
1071 AGAGTTGATATATTGCGCAGAGCTGGATGAGCTCCTGAGGGAATAATGCTGGTCCAT 1130  
181 CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACCTGAATATTAT 240  
1131 CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACCTGAATATTAT 1190  
241 CCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATCTGT 300  
1191 CCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATCTGT 1250  
301 GAGCCACTAATATTCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 360  
1251 GAGCCACTAATATTCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 1310  
361 TCATGTTTTTCCCAAGTCACGAAAGAGAAATTTGAGTTTATTTTTGCTCTGAAATGCAA 420  
1311 TCATGTTTTTCCCAAGTCACGAAAGAGAAATTTGAGTTTATTTTTGCTCTGAAATGCAA 1370  
421 AACAGGTTTTCCGTCATTTATACAAATTTACATCTATTATTTAAAGGAAGCAATATAACG 480  
1371 AACAGGTTTTCCGTCATTTATACAAATTTACATCTATTATTTAAAGGAAGCAATATAACG 1430  
481 ATCCAGTGTGGGCTCGCTGCCCAAGTGATTTTCAAGTGTCTTATCAAGAGAGATAGC 540  
1431 ATCCAGTGTGGGCTCGCTGCCCAAGTGATTTTCAAGTGTCTTATCAAGAGAGATAGC 1490  
541 AATTACAGTGTGATGGAGTTCCTTGGCCGCGATGATCTAATATCCAGTTGATGA 600  
1491 AATTACAGTGTGATGGAGTTCCTTGGCCGCGATGATCTAATATCCAGTTGATGA 1550  
601 AGTCAGAGGCTGTATATTGAGGCAACCAAGACTCCCTTTAGAGCATCACTGTA 660  
1551 AGTCAGAGGCTGTATATTGAGGCAACCAAGACTCCCTTTAGAGCATCACTGTA 1610  
661 CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA 720  
1611 CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA 1670  
721 TTCTTGCTGCTACAGTCAGCTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAGAA 780  
1671 TTCTTGCTGCTACAGTCAGCTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAGAA 1730  
781 TCACACTGTGTGTCCCTTTTAAAGCTATCAAGTCTCTGAAAGATGACCCCACTTGCAAAAC 840  
1731 TCACACTGTGTGTCCCTTTTAAAGCTATCAAGTCTCTGAAAGATGACCCCACTTGCAAAAC 1790  
841 AAAGGAATTTTGGGCGCACATTTTGGATTTCAGCAGGTCCTCTTCTCTGACTACTCTCC 900  
1791 AAAGGAATTTTGGGCGCACATTTTGGATTTCAGCAGGTCCTCTTCTCTGACTACTCTCC 1850  
901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGTATGGGATGCTTACAGGCC 960  
1851 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGTATGGGATGCTTACAGGCC 1910  
961 TCATGATCTACAGCTGGAAGAAATATCTCTGCTGCTCTCATATATGTTGGTCTTCA 1020  
1911 TCATGATCTACAGCTGGAAGAAATATCTCTGCTGCTCTCATATATGTTGGTCTTCA 1970  
1021 GGTTCGCTATGCTGGGCGCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACCGA 1080  
1971 GGTTCGCTATGCTGGGCGCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACCGA 2030  
1081 AGTTTATAGGTCACCTGACCAAGATGACAGGGCTATTTACTAGGATCTGTGGCCAT 1140  
2031 AGTTTATAGGTCACCTGACCAAGATGACAGGGCTATTTACTAGGATCTGTGGCCAT 2090

QY 1141 GCAGCAGCAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200  
Db 2091 GCAGCAGCAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 2150  
QY 1201 TGAGATGTCATTTTGGACATACACAGATATATTACTGAGTGTGTGTAGTGAGGGCTGGA 1260  
Db 2151 TGAGATGTCATTTTGGACATACACAGATATATTACTGAGTGTGTGTAGTGAGGGCTGGA 2210  
QY 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAAGTTCCTGAATCGGG 1320  
Db 2211 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAAGTTCCTGAATCGGG 2270  
QY 1321 AGAACATTTATGAACATGCTCTTTTGCACCTACCTTCAAGAAAACTTGGATCACGATATGC 1380  
Db 2271 AGAACATTTATGAACATGCTCTTTTGCACCTACCTTCAAGAAAACTTGGATCACGATATGC 2330  
QY 1381 TGCTCTAAAAGTGATATA 1399  
Db 2331 TGCTCTAAAAGTGATATA 2349  
RESULT 10  
AAH99934  
ID AAH99934 standard; cDNA; 3143 Bp.  
XX AC AAH99934;  
XX AC AAH99934;  
XX 12-APR-2002 (first entry)  
XX cDNA encoding 21953 human prollyl oligopeptidase.  
XX 21953 prollyl oligopeptidase; human; proline; endopeptidase; cancer;  
XX cardiovascular disease; autoimmune disease; atopic allergy;  
XX neurological disorder; vascular disorder; prostate disorder; cycostatic;  
XX antidiabetic; antichratic; antidiabetic; antiinflammatory;  
XX diabetes mellitus; arthritis; multiple sclerosis; asthma;  
XX Grave's disease; neuronal disorder; demyelinating disease; ss.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX COS 229.2877  
XX /tag= a  
XX /product= "21953 prollyl oligopeptidase"  
XX /note= "This region is specifically claimed in claim 2"  
WO200179473-A2.  
XX 25-OCT-2001.  
XX 11-APR-2001; 2001WO-US040483.  
XX 18-APR-2000; 2000US-0197508P.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Meyers RA, Williamson M;  
XX WPI; 2002-034353/04.  
XX P-PSDB; AAG78415.  
XX New polypeptides 21953, member of human prollyl oligopeptidase family,  
XX useful as diagnostic targets and therapeutic agents for controlling  
XX cancer, lymphoma and leukemia.  
XX Claim 7; Page 100-102; 121pp; English.  
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.  
XX which is cytosolic, antidiabetic, antarthritic, neuroprotective,  
XX antithyroid, dermatological, antipsoriatic, antidiabetic,  
XX ophthalmological, antiinflammatory, neurotropic, antiparkinsonian,  
XX anticonvulsant, gynaecological, vasotropic, cardiant,

CC antiatherosclerotic, anorectic and metabolic in its action. Uses include  
CC gene therapy, expression or activity of 21953 protein modulator, it is  
CC useful for identifying a compound which binds to it and can be used in  
CC preventing, treating or detecting a cellular proliferative or  
CC differentiative disorder. The 21953 molecules can act as novel diagnostic  
CC targets and therapeutic agents for controlling disorders associated with  
CC the aberrant activity or degradation of peptide hormones e.g., disorders  
CC associated with cell differentiation and proliferation such as cancer,  
CC immune function, reproductive, neurological and cardiovascular function.  
CC The 21953 molecules are thus useful for treating and preventing cellular  
CC proliferative and differentiative disorders, haematopoietic neoplastic  
CC disorders, immune disorders such as autoimmune diseases, diabetes  
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,  
CC neuronal disorders, demyelinating diseases, vascular disorders and  
CC metabolism or pain disorders. This sequence represents the cDNA encoding  
CC sequence of 21953 human prolyl oligopeptidase  
XX  
SQ Sequence 3143 BP; 943 A; 644 C; 712 G; 844 T; 0 U; 0 Other;  
Query Match 81.1%; Score 1353.4; DB 6; Length 3143;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 1664; Conservative 0; Mismatches 1; Indels 300; Gaps 1;  
QY 1 AACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAGATGTCAGAAATTAATGATTGATGCTGA 60  
DB 1179 AACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAGATGTCAGAAATTAATGATTGATGCTGA 1238  
QY 61 AGGAGGATCATAGATGTCATAGATAGGAACTTAATCAACCTTTTGAGATCTTATTGA 120  
DB 1239 AGGAGGATCATAGATGTCATAGATAGGAACTTAATCAACCTTTTGAGATCTTATTGA 1298  
QY 121 AGGAGTGAATATATATGCGAGCTGGATGACTCTCTGAGGAGAAATATGCTTGGTCCAT 180  
DB 1299 AGGAGTGAATATATATGCGAGCTGGATGACTCTCTGAGGAGAAATATGCTTGGTCCAT 1358  
QY 181 CCTACTAGATCGCTCCAGACTGCGCTACAGATAGTGTGATCTCACCTGGAATATTAT 240  
DB 1359 CCTACTAGATCGCTCCAGACTGCGCTGCGATAGTGTGATCTCACCTGGAATATTAT 1418  
QY 241 CCCAGTAGAAGATGATGTTATGGAAGCGACAGACTCATTTGAGTCAGTGGCTGATTCGT 300  
DB 1419 CCCAGTAGAAGATGATGTTATGGAAGCGACAGACTCATTTGAGTCAGTGGCTGATTCGT 1478  
QY 301 GAGGCCACTAAATATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360  
DB 1479 GAGGCCACTAAATATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1538  
QY 361 TCATGTTTTTCCCAAGTCACGAGAGGAAATTTGATTTATTTTTTGGCTCTCGAATGCAA 420  
DB 1539 TCATGTTTTTCCCAAGTCACGAGAGGAAATTTGATTTATTTTTTGGCTCTCGAATGCAA 1598  
QY 421 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACG 480  
DB 1599 AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACG 1658  
QY 481 ATCCAGTGTGGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGC 540  
DB 1659 ATCCAGTGTGGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGC 1718  
QY 541 AATTAACAGTGTGGAATGGGAAGTTCTTGGCGGGATGGATCTAATATCAAGTTGATGA 600  
DB 1719 AATTAACAGTGTGGAATGGGAAGTTCTTGGCGGGATGGATCTAATATCAAGTTGATGA 1778  
QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTACAGCATCACCTGTA 660  
DB 1779 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTACAGCATCACCTGTA 1838  
QY 661 CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA 720  
DB 1839 CGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA 1898  
QY 721 TTCTTGTGATCAGTCAGACATGTGACTTTTATTAAGTAAGTATATAGTAAACGAGAA 780

DB 1899 TTCTTGTGATCAGTCAGACCTGTGACTTCTTTTAAAGTAAGTATAGTAAACGAGAA 1958  
QY 781 TCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGC AAAAC 840  
DB 1959 TCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGC AAAAC 2018  
QY 841 AAAGGAATTTGGGCCACCATTTTGGATTGAGAGTCCCTCTTCTGACTATACTCTCC 900  
DB 2019 AAAGGAATTTGGGCCACCATTTTGGATTGAGAGTCCCTCTTCTGACTATACTCTCC 2078  
QY 901 AGAAATTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGGATGCTCTTACAAGCC 960  
DB 2079 AGAAATTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGGATGCTCTTACAAGCC 2138  
QY 961 TCATGATCTACAGCCTGGAAGAAATATCTCTACTGTCTCTTCAATATATGATGGTGGTCTCA 1020  
DB 2139 TCATGATCTACAGCCTGGAAGAAATATCTCTACTGTCTCTTCAATATATGATGGTGGTCTCA 2198  
QY 1021 ----- 1020  
DB 2199 GGTGCAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCTAGC 2258  
QY 1021 ----- 1020  
DB 2259 CTCTCTAGGTTATGTGTGTAGTATAGACAAACAGGGGATCCTGTCAACGAGGGCTTAA 2318  
QY 1021 ----- 1020  
DB 2319 ATTTGAAGCGCCTTTTAAATATATAAATGdGTCAAAATAGAAAATTGACGATCAGGTGGAAGG 2378  
QY 1021 ----- 1020  
DB 2379 ACTCCAATATCTAGCTTCTCGATATGATTTCAATTGACTTGTGGGCTATCCACGG 2438  
QY 1021 ----- 1020  
DB 2439 CTGGTCTATGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTCAGATATCTTCAG 2498  
QY 1021 GGTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGATCTTTATGATACAGGATACACGGA 1080  
DB 2499 GGTGTCTATTTGCTGGGGCCCCAGTCACTCTGTGATCTTTATGATACAGGATACACGGA 2558  
QY 1081 ACCTTATATGGGTCACTCCCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140  
DB 2559 ACCTTATATGGGTCACTCCCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCAT 2618  
QY 1141 GCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGATGTTTCTTGA 1200  
DB 2619 GCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGATGTTTCTTGA 2678  
QY 1201 TGAAATGTCCATTTTGGCAATATATTAATCTGATGATTTTGTAGTGGGCTGAAA 1260  
DB 2679 TGAAATGTCCATTTTGGCAATATATTAATCTGATGATTTTGTAGTGGGCTGAAA 2738  
QY 1261 GCATATGATTTACAGATCTATCTCCAGAGAGACACAGATAGAGTTCTCTGAATCGGG 1320  
DB 2739 GCATATGATTTACAGATCTATCTCCAGAGAGACACAGATAGAGTTCTCTGAATCGGG 2798  
QY 1321 AGAACATTTATGAACCTCATCTTTTGCATCTACCTTCAAGAAAAACCTTGGATCACTATTGC 1380  
DB 2799 AGAACATTTATGAACCTCATCTTTTGCATCTACCTTCAAGAAAAACCTTGGATCACTATTGC 2858  
QY 1381 TGCTCTAAAAGTGAATATAATTTTGGATCTCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440  
DB 2859 TGCTCTAAAAGTGAATATAATTTTGGATCTCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2918  
QY 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATTTTGAATACC 1500  
DB 2919 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATTTTGAATACC 2978  
QY 1501 TGCCATGTAACTACTCTCTGAAAAATAAATGTTGGTGGCCATGACAGGGGTCTACGGTTGT 1560  
DB 2979 TGCCATGTAACTACTCTCTGAAAAATAAATGTTGGTGGCCATGACAGGGGTCTACGGTTGT 3038



Db 830 GCGCCTTTAAATATATAAATGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCAA 771  
Qy 1021 ----- 1020  
Db 770 TAATAGCTTCTCGATATGATTTCAATGACTTAGATCGTGTGGCATCCACGGCTGTGTC 711  
Qy 1021 -----GTTTGT 1027  
Db 710 TATGGAGTACCTCTCCCTGATGCGATTAATGACAGGTCAGATATCTTCAAGGTTGCT 651  
Qy 1028 ATTGCTGGGCCCCAGTCACTCTGTGGATCTTATGATACAGGATACAGGAACGTTAT 1087  
Db 650 ATTGCTGGGCCCCAGTCACTCTGTGGATCTTATGATACAGGATACAGGAACGTTAT 591  
Qy 1088 ATGGGTCACTTACAGCAATGAACAGGGCTATTACTTAGGATCTGTGGCATCAAGCA 1147  
Db 590 ATGGGTCACTTACAGCAATGAACAGGGCTATTACTTAGGATCTGTGGCATCAAGCA 531  
Qy 1148 GAAAGTTCCCTCTGAAACCAATGTTTACTGCTCTTACATGTTTCTTGGATGAGAT 1207  
Db 530 GAAAGTTCCCTCTGAAACCAATGTTTACTGCTCTTACATGTTTCTTGGATGAGAT 471  
Qy 1208 GTCCATTTTGCACATACAGTATATTACTAGTGTGAGGGCTGGAAAGCCATAT 1267  
Db 470 GTCCATTTTGCACATACAGTATATTACTAGTGTGAGGGCTGGAAAGCCATAT 411  
Qy 1268 GATTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCCTGAATCGGAGAACAT 1327  
Db 410 GATTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCCTGAATCGGAGAACAT 351  
Qy 1328 TATGAAGTCACTTTTGGACTGCTGAGAAACCTTGAAGAAACCTTGGATCAGTATTGCTCTA 1387  
Db 350 TATGAAGTCACTTTTGGACTGCTGAGAAACCTTGAAGAAACCTTGGATCAGTATTGCTCTA 291  
Qy 1388 AAGTGTATATATTTTGAAGTCTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAA 1447  
Db 290 AAGTGTATATATTTTGAAGTCTGTAGAACTCTCTGGTATACACTGGCTATTTAAACCAA 231  
Qy 1448 TGAGGAGTTTAAATCAACAGAAACACAGAAATTTGATCATCATATTTTATACCTGCGCATG 1507  
Db 230 TGAGGAGTTTAAATCAACAGAAACACAGAAATTTGATCATCATATTTTATACCTGCGCATG 171  
Qy 1508 TATCATCTACTCTGAAATAAATGTGTGCGCATGACAGGGTCTACGGTTTGTGTAGTA 1567  
Db 170 TATCATCTACTCTGAAATAAATGTGTGCGCATGACAGGGTCTACGGTTTGTGTAGTA 111  
Qy 1568 ATCTAATACCTTAACCCCATGCTCAAAATCAAAATGATATATATTTCTGAGAGCCAG 1627  
Db 110 ATCTAATACCTTAACCCCATGCTCAAAATCAAAATGATATATATTTCTGAGAGCCAG 51  
Qy 1628 CAATACATAGAAATTTACTTAAAAAATAAAAAAAAAAAAAAAAAAAAA 1667  
Db 50 CAATACATAGAAATTTACTTAAAAAATAAAAAAAAAAAAAAAAAAAAA 11

RESULT 12  
ABK83327  
ID ABK83327 standard; cDNA; 4829 BP.  
XX  
AC ABK83327;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
XX cDNA encoding human DPRP-1 splice variant #3.  
DE  
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;  
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;  
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;  
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;  
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;  
KW dyskinesia; reproductive disorder; inflammatory disorder;  
XX metabolic disorder; gene; ss.  
XX

OS Homo sapiens.  
XX WO200231134-A2.  
XX PD 18-APR-2002.  
XX PF 12-OCT-2001; 2001WO-US031874.  
XX PR 12-OCT-2000; 2000US-0240117P.  
XX (FERR ) FERRING BV.  
PA Qi S, Akinsanya KO, Riviere PJ, Junien J;  
PI WPI; 2002-444178/47.  
XX P-PSDB; ABG61596.  
DR New dipeptidyl peptidase IV-related proteins and nucleic acids encoding  
XX the proteins, useful for treating e.g. fungal, bacterial, protozoan and  
PT viral infections, cancers, allergies, neurological disorders, or pain.  
PT  
XX Disclosure; Page 65-66; 113pp; English.  
XX  
CC The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins  
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic  
CC acids encoding them are useful for treating infections such as fungal,  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322  
CC -ABK83343 encode human DPRP proteins  
XX  
SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;  
Query Match 80.9%; Score 1349.6; DB 6; Length 4829;  
Best Local Similarity 84.5%; Pred. No. 0;  
Matches 1664; Conservative 0; Mismatches 4; Indels 302; Gaps 1;  
Qy 1 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTGAGAAATTAATGATGCTGA 60  
Db 1164 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTGAGAAATTAATGATGCTGA 1223  
Qy 61 AGGAAGATCATAGATGTGATAGATAGAACTAAATTCACCTTTTGAGATTCATTTGA 120  
Db 1224 AGGAAGATCATAGATGTGATAGATAGAACTAAATTCACCTTTTGAGATTCATTTGA 1283  
Qy 121 AGGAGTTGATATATTTGCCAGAGCTGGATGCTCTCGAGGAAATATGCTGTCAT 180  
Db 1284 AGGAGTTGATATATTTGCCAGAGCTGGATGCTCTCGAGGAAATATGCTGTCAT 1343  
Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 240  
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT 1403  
Qy 241 CCCAGTAGAAGATGATGTTATGGAAAGGACAGAGATCAATGAGTCAGTGCCTGATTCGT 300  
Db 1404 CCCAGTAGAAGATGATGTTATGGAAAGGACAGAGATCAATGAGTCAGTGCCTGATTCGT 1463  
Qy 301 GAGCCACTAATTTATCTATGAGAAACACAGACATCTGGATTAATATCCATCATCTT 360  
Db 1464 GAGCCACTAATTTATCTATGAGAAACACAGACATCTGGATTAATATCCATCATCTT 1523  
Qy 361 TCATGTTTTTCCCAAGTCAACAGAGGAAATTCAGTTTATTTTGGCTCTGAATGCAA 420  
Db 1524 TCATGTTTTTCCCAAGTCAACAGAGGAAATTCAGTTTATTTTGGCTCTGAATGCAA 1583  
Qy 421 AACAGGTTTCCGTCATTTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAACG 480

Db	1584	 AACAGTTCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATAATAACG	1643
Qy	481	ATCCAGTGGTGGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAAGAGGAGATAGC	540
Db	1644	ATCCAGTGGTGGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAAGAGGAGATAGC	1703
Qy	541	AATTACCAGTGGTGAATGGGAAGTTCCTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	600
Db	1704	AATTACCAGTGGTGAATGGGAAGTTCCTGGCCGGCATGGATCTAATATCCAAAGTTGATGA	1763
Qy	601	AGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAAGACTCCCTTTTAGAGCATCACCTGTGA	660
Db	1764	AGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAAGACTCCCTTTTAGAGCATCACCTGTGA	1823
Qy	661	CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTCACA	720
Db	1824	CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACCAAGGCTGACTGACCGTGGCTACTCACA	1883
Qy	721	TTCTGTCGTCATCAGTCAGACATGTGACTTCCTTTAATAGTAAGTATAGTACACAGAGAA	780
Db	1884	TTCTGTCGTCATCAGTCAGACATGTGACTTCCTTTAATAGTAAGTATAGTACACAGAGAA	1943
Qy	781	TCCACATCTGTGTCCCTTTACAGCTATCAAGTCTCGAAGTGCCTGAGTAGCCCACTTCGCAAC	840
Db	1944	TCCACATCTGTGTCCCTTTACAGCTATCAAGTCTCGAAGTGCCTGAGTAGCCCACTTCGCAAC	2003
Qy	841	AAAGGAATTTTGGGCGCACCAATTTTGGANTCAGCAGGTCCTCTTCTGACTATACTCTCTCC	900
Db	2004	AAAGGAATTTTGGGCGCACCAATTTTGGANTCAGCAGGTCCTCTTCTGACTATACTCTCTCC	2063
Qy	901	AGAAATTTCTCTTTTGAAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCC	960
Db	2064	AGAAATTTCTCTTTTGAAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCC	2123
Qy	961	TCATGATCTACAGCCTGGAAAGAAATACCTACTGTGCTGTTTCATATATGGTGGTCTCTCA	1020
Db	2124	TCATGATCTACAGCCTGGAAAGAAATACCTACTGTGCTGTTTCATATATGGTGGTCTCTCT	2183
Qy	1021	-----	1020
Db	2184	CAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGATATTTCCGCTTGAATACCCCTA	2243
Qy	1021	-----	1020
Db	2244	GCCTCTCTAGTTATGTGTTGTAGTAGATAGACAACAGGGGATCCTGTACCAGGGGCTT	2303
Qy	1021	-----	1020
Db	2304	AAATTTGAAGGCGCTTTAAATATAAATGGGTCAATAGAAATTGACGATCAGGTGGAA	2363
Qy	1021	-----	1020
Db	2364	GGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGATCTGTAGATCTGTGGGCATCCAC	2423
Qy	1021	-----	1020
Db	2424	GGCTGGTCCTATGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTC	2483
Qy	1021	--GGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGC	1078
Db	2484	AGGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGC	2543
Qy	1079	GAACTGTTATGGGTCACCTCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCC	1138
Db	2544	GAACTGTTATGGGTCACCTCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCC	2603
Qy	1139	ATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTCTTACATGGTTTCTTG	1198
Db	2604	ATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTCTTACATGGTTTCTTG	2663
Qy	1199	GATGAGAAATGCTCAATTTTGACATACAGTATATTACTGAGTTTTTTTAGTGAGGGCTGGA	1258

RESULT 13	
AAC85694	
ID	AAC85694 standard; cDNA; 3120 BP.
XX	
XX	AAC85694;
XX	
DT	29-JUN-2001 (first entry)
XX	
DE	Nucleotide sequence of human DPP8.
XX	
KW	Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase; dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea; growth hormone deficiency; glucose level; mucosal regeneration; non-insulin dependent diabetes mellitus; glucose intolerance; immunosuppression; es.
KW	
KW	
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
PH	214..2862
CDS	/*tag= a
FT	/product= "Human DPP8"
FT	
XX	
PN	WO200119866-A1.
PD	
XX	22-MAR-2001.
XX	
PF	11-SEP-2000; 2000WO-AU001085.
XX	
PR	10-SEP-1999; 99AU-00002762.
PR	18-FEB-2000; 2000AU-00005709.
XX	
PA	(UNSY ) UNIV SYDNEY.
XX	
PI	Abbott CA, Gorell MD;
XX	
DR	WPI; 2001-281520/29.
DR	P-PSDB; AAB47187.
XX	
PT	New human dipeptidyl aminopeptidase (DPP8) useful for cleaving substrates, identifying inhibitors of DPP8 catalytic activity wh
PT	therapeutic uses and for detecting activated T cells.









Db 2304 ATTGGAAGGCGCTTTAAATATAAAATGGGTCAAAATAGAAATTCACGATCAGTGAAGG 2363  
Qy 1021 ----- 1020  
Db 2364 AC7CCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATCCACGG 2423  
Qy 1021 ----- 1020  
Db 2424 CTGGTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAG 2483  
Qy 1021 GGTTCGTATTGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 1080  
Db 2484 GGTTCGTATTGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA 2543  
Qy 1081 ACCTTATATGGGTCACTCTGACCAAGAAATGAACAGGGCTATTACTAGGATCTGTGGCCAT 1140  
Db 2544 ACCTTATATGGGTCACTCTGACCAAGAAATGAACAGGGCTATTACTAGGATCTGTGGCCAT 2603  
Qy 1141 GCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGA 1200  
Db 2604 GCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTTGA 2663  
Qy 1201 TGGATATGTCATTTTGGACATACAGTATATTACTGAGTTTCTTACTGAGGGCTGAAA 1260  
Db 2664 TGGATATGTCATTTTGGACATACAGTATATTACTGAGTTTCTTACTGAGGGCTGAAA 2723  
Qy 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATAAGAGTTCTCGAATCGGG 1320  
Db 2724 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATAAGAGTTCTCGAATCGGG 2783  
Qy 1321 AGAACATTTAGAACTGCATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGTATGTC 1380  
Db 2784 AGAACATTTAGAACTGCATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGTATGTC 2843  
Qy 1381 TGCTCTAAAGTGATATATTTTGGACCTGTGTGAGAACTCTCTGTATACACTGCTATTT 1440  
Db 2844 TGCTCTAAAGTGATATATTTTGGACCTGTGTGAGAACTCTCTGTATACACTGCTATTT 2903  
Qy 1441 AACCAAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATCTTTGTATACC 1500  
Db 2904 AACCAAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATCTTTGTATACC 2963  
Qy 1501 TGCCATGTAACTACTCTCTGAAATAAATGTTGTGTCATGCGGGGTCTACGGTTTGT 1560  
Db 2964 TGCCATGTAACTACTCTCTGAAATAAATGTTGTGTCATGCGGGGTCTACGGTTTGT 3023  
Qy 1561 GGTAGTAACTATACCTTAAACCCACATGCTCAAAATCAAAATGATCATCATCTTTCTGAGA 1620  
Db 3024 GGTAGTAACTATACCTTAAACCCACATGCTCAAAATCAAAATGATCATCATCTTTCTGAGA 3083  
Qy 1621 GACCCAGCAATACCATAAGAAATTTACTAAAAA 1657  
Db 3084 GACCCAGCAATACCATAAGAAATTTACTAAAAA 3120

RESULT 15  
ABK12892  
ID ABK12892 standard; cDNA; 3106 BP.  
XX  
AC  
XX  
DT  
XX  
DE  
XX  
KW  
KW  
KW  
KW  
KW  
KW  
OS  
Human protease PRTS-9 cDNA sequence.  
Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
cell proliferative disorder; developmental disorder; epilepsy;  
Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
reproductive disorder; endometriosis; ss.  
Homo sapiens.

Query Match 80.5%; Score 1344; DB 6; Length 3106;  
Best Local Similarity 84.6%; Pred. No. 0;  
Matches 1654; Conservative 0; Mismatches 0; Indels 300; Gaps 1;  
Qy 1 AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTGAGAAATPAATGATGATGCTGA 60  
Db 1153 AACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTGAGAAATPAATGATGCTGA 1212  
Qy 61 AGGAAGATCATAGATGTCTATAGATAGGAACTTAATTCACCTTTTGAGATCTTATTTGA 120  
Db 1213 AGGAAGATCATAGATGTCTATAGATAGGAACTTAATTCACCTTTTGAGATCTTATTTGA 1272  
Qy 121 AGGAGTTGAATATATTGCGAGAGCTGGATGAGTCTCTGAGGAAATATGCTTGTGTCAT 180  
Db 1273 AGGAGTTGAATATATTGCGAGAGCTGGATGAGTCTCTGAGGAAATATGCTTGTGTCAT 1332  
Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCCTGATGATTTATTTAT 240  
Db 1333 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCCTGATGATTTATTTAT 1392  
Qy 241 CCCAGTGAAGATGATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTCTGATCTTGT 300

XX Key Location/Qualifiers  
FH CDS 203..2851  
FT /\*tag= a  
FT /product= "Human protease PRTS-9"  
XX  
FN WO200198468-A2.  
XX  
PD 27-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019178.  
XX  
PR 16-JUN-2000; 2000US-0212336P.  
XX  
PR 22-JUN-2000; 2000US-0213955P.  
XX  
PR 29-JUN-2000; 2000US-0215396P.  
XX  
PR 07-JUL-2000; 2000US-0216821P.  
XX  
PR 14-JUL-2000; 2000US-0218946P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX

PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;  
PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;  
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
PI Kallick DA;  
XX  
DR WP1; 2002-090437/12.  
XX P-PSDB; AAU74749.  
XX  
PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in  
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative  
PT (e.g. cancer) disorders.  
XX  
PS Claim 5; Page 166-167; 177pp; English.  
XX  
CC The present invention relates to twenty one new human proteases, referred  
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the  
CC invention are useful in the diagnosis, treatment and prevention of  
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's  
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial  
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency  
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.  
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,  
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's  
CC disease and reproductive e.g. infertility and endometriosis disorders.  
CC Numerous other examples of each disorder are given in the specification.  
CC The present nucleic acid sequence encodes the human protease PRTS-9  
CC protein of the invention  
XX  
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:59:51 ; Search time 6252.09 Seconds  
(without alignments)  
12489.839 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 1669  
Sequence: 1 aacaggtaacgaatccta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_est4: \*  
5: gb\_est5: \*  
6: gb\_est6: \*  
7: gb\_est7: \*  
8: gb\_gss1: \*  
9: gb\_gss2: \*  
10: gb\_gss3: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1568	93.9	2292	4 CR609512	full-length cDNA clone CSODL005YD02 of B cells (Ramos cell line)
2	904.6	54.2	3143	4 AK016546	Mus muscu
3	888.4	53.2	2649	10 AY411615	AY411615 Homo sapi
4	879.4	52.7	5517	4 AK029788	Mus muscu
5	842.8	50.5	2649	10 AY411616	AY411616 Pan trogl
6	820.6	49.2	910	5 BX372276	BX372276 BX372276
7	804.4	48.2	902	8 DN517062	DN517062 1257572 M
8	730.6	43.8	857	8 DN519771	DN519771 1262448 M
9	688.4	41.2	853	2 BI223892	BI223892 602941035
10	675.6	40.5	2634	10 AY411617	AY411617 Mus muscu
11	672.6	40.3	1013	7 CN647232	CN647232 ILLUMIGEN
12	672.6	40.3	1027	7 CN646612	CN646612 ILLUMIGEN
13	660	39.5	993	2 BG259714	BG259714 602380072
14	646.4	38.7	1292	4 AF175225	AF175225 Homo sapi
15	644	38.6	722	5 BU631054	BU631054 UI-R-Fel-
16	640.8	38.4	804	5 BU687687	BU687687 UI-CF-ECL
17	637	38.2	1038	1 AL582206	AL582206 AL582206
18	628	37.6	634	1 AI819365	AI819365 wg61c02.x
19	623.6	37.4	760	1 AI917735	AI917735 tt1ld07.x
20	622.8	37.3	673	1 AW303607	AW303607 xv21b12.x
21	620	37.1	864	1 AL542617	AL542617 AL542617
22	607.6	36.4	1278	4 AF173382	AF173382 Homo sapi

#### ALIGNMENTS

RESULT 1  
LOCUS CR609512 2292 bp mRNA linear HTC 21-JUL-2004  
DEFINITION full-length cDNA clone CSODL005YD02 of B cells (Ramos cell line)  
Cot 25-normalized of Homo sapiens (human).  
ACCESSION CR609512  
VERSION CR609512.1 GI:50490319  
KEYWORDS HTC; CNSLT\_CDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2292)  
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
Genoscope.  
REFERENCE 2 (bases 1 to 2292)  
AUTHORS Direct Submission  
JOURNAL Genoscope.  
SUBMITTED (20-JUL-2004) Genoscope - Centre National de Sequençage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr )  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
FEATURES  
source Location/Qualifiers  
1..2292  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODL005YD02"  
/tissue\_type="B cells (Ramos cell line) Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"

#### ORIGIN

Query Match 93.9%; Score 1568; DB 4; Length 2292;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1590; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

QY 1 AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGTGCAGAAATAATGATGATGCTGA 60  
Db 703 AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGTGCAGAAATAATGATGATGCTGA 762  
QY 61 AGGAAGGATCATAGATGTCATAGATAAGGAATCAATCAACTTTTGAAGATTCATTTGA 120  
Db 763 AGGAAGGATCATAGATGTCATAGATAAGGAATCAATCAACTTTTGAAGATTCATTTGA 822  
QY 121 AGGAGTTGAATATATATGCCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCAT 180  
Db 823 AGGAGTTGAATATATATGCCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCAT 882  
QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATATTTAT 240  
Db 883 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATATTTAT 942  
QY 241 CCCAGTAGAAGATGATGTTATGGAAGGAGAGAGACTCAATGAGTCAGTGCCTGATTCCTGT 300  
Db 943 CCCAGTAGAAGATGATGTTATGGAAGGAGAGAGACTCAATGAGTCAGTGCCTGATTCCTGT 1002  
QY 301 GAGCCCACTAATATCTATGAGAACACACACATCTGGATAATATCCATGACATCTT 360  
Db 1003 GAGCCCACTAATATCTATGAGAACACACACATCTGGATAATATCCATGACATCTT 1062  
QY 361 TCATGTTTTCCTCCAAAGTCACGAAGAGGAATTCAGTTTATTTTGGCTCTGAATGCAA 420  
Db 1063 TCATGTTTTCCTCCAAAGTCACGAAGAGGAATTCAGTTTATTTTGGCTCTGAATGCAA 1122  
QY 421 AACAGGTTTCGCTCATTTATACAAATTAATCATCTATTTTAAAGGAAAGCAATATAACG 480  
Db 1123 AACAGGTTTCGCTCATTTATACAAATTAATCATCTATTTTAAAGGAAAGCAATATAACG 1182  
QY 481 ATCCAGTGTGGGTGCTGCTCCAAAGTGAATTCAGTGTCTTATCAAGGTGCTATCAAGAGGAGATGC 540  
Db 1183 ATCCAGTGTGGGTGCTGCTCCAAAGTGAATTCAGTGTCTTATCAAGGTGCTATCAAGAGGAGATGC 1242  
QY 541 AATTACCAGTGGTGAATGGGAAGTCTTGGCCGGCATGATCTAATATCAAGTTGATGA 600  
Db 1243 AATTACCAGTGGTGAATGGGAAGTCTTGGCCGGCATGATCTAATATCAAGTTGATGA 1302  
QY 601 AGTCAGAAAGGCTGTGTATATTTGAAGGCCCAAGAGCTCCCTTTAGAGCATCACTCTGTA 660  
Db 1303 AGTCAGAAAGGCTGTGTATATTTGAAGGCCCAAGAGCTCCCTTTAGAGCATCACTCTGTA 1362  
QY 661 CGTAGTCAGTTAAGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGCTACTCACA 720  
Db 1363 CGTAGTCAGTTAAGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGCTACTCACA 1422  
QY 721 TTCTTGTGTCATCAGTCAGTCAGTCAGTCTTCTTTATAGTAAAGTATAGTAAACCAAGAA 780  
Db 1423 TTCTTGTGTCATCAGTCAGTCAGTCAGTCTTCTTTATAGTAAAGTATAGTAAACCAAGAA 1482  
QY 781 TCCACACTGTGTCTCCTTTACAAAGTATCAAGTCTCTGAAGTACCCCAACTTCCAAAC 840  
Db 1483 TCCACACTGTGTCTCCTTTACAAAGTATCAAGTCTCTGAAGTACCCCAACTTCCAAAC 1542  
QY 841 AAAGGAATTTTGGCCACATTTTGGATTCAGCAGGTCTCTTCTGACTACTACTCTCC 900  
Db 1543 AAAGGAATTTTGGCCACATTTTGGATTCAGCAGGTCTCTTCTGACTACTACTCTCC 1602  
QY 901 AGAAATTTCTCTTTGAAGTACTACTGATTTACATGTTAGGAGTCTCTCAAGCC 960  
Db 1603 AGAAATTTCTCTTTGAAGTACTACTGATTTACATGTTAGGAGTCTCTCAAGCC 1662  
QY 961 TCATGATCTACAGCTGGAAGAAATATCTCTACTGTGCTGTTTATATATGTTGCTCTCA 1020  
Db 1663 TCATGATCTACAGCTGGAAGAAATATCTCTACTGTGCTGTTTATATATGTTGCTCTCA 1722  
QY 1021 GGTGTCTATGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGA 1080  
Db 1723 GGTGTCTATGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGA 1782  
QY 1081 ACGTTATATGGGTCAACCTGACCAAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGCCAT 1140

Db 1783 ACGTTATATGGGTCAACCTGACCAAGAAATGAACAGGGCTATTACTTTAGATCTGTGGCCAT 1842  
QY 1141 GCAAGCAGAAAAGTTCCTCTCGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200  
Db 1843 GCAAGCAGAAAAGTTCCTCTCGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1902  
QY 1201 TGAGAATGTCATTTTGGCAGATACCAAGTATATTTACTGAGTGTGAGGGCTGGAAA 1260  
Db 1903 TGAGAATGTCATTTTGGCAGATACCAAGTATATTTACTGAGTGTGAGGGCTGGAAA 1962  
QY 1261 GCCATATGATTTTACAGATCTATCCTCAGGAGAGACACAGCATTAAGAGTTCCTGAATCGGG 1320  
Db 1963 GCCATATGATTTT-----CAGGAGAGACACAGCATTAAGAGTTCCTGAATCGGG 2010  
QY 1321 AGAACATTTATGAATCGCATCTTTTGGCAGTACCTTCAAGAAAAACCTTGGATCACGTATTGC 1380  
Db 2011 AGAACATTTATGAATCGCATCTTTTGGCAGTACCTTCAAGAAAAACCTTGGATCACGTATTGC 2070  
QY 1381 TGCTCTAAAAGTGTATATAATTTTGGCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440  
Db 2071 TGCTCTAAAAGTGTATATAATTTTGGCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2130  
QY 1441 AACCAATGAGGAGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATTTTGTATACC 1500  
Db 2131 AACCAATGAGGAGTTTAAATCAACAGAAAAACACAGAAATTTGATCATCATTTTGTATACC 2190  
QY 1501 TGCCATGTAAATCTACTCTGAAATTAATCTGTTGTCATGCAGGGCTTACCGTTTGT 1560  
Db 2191 TGCCATGTAAATCTACTCTGAAATTAATCTGTTGTCATGCAGGGCTTACCGTTTGT 2250  
QY 1561 GGTAGTAAATCTATACCTTTAAACCCAGCATGCTCAAAATCAA 1602  
Db 2251 GGTAGTAAATCTATACCTTTAAACCCAGCATGCTCAAAATCAA 2292

RESULT 2  
AKO16546 3143 bp mRNA linear HTC 03-APR-2004  
LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932434F09 product:DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens], full insert sequence.  
DEFINITION AKO16546  
ACCESSION AKO16546.1 GI:12855334  
VERSION HTC; CAP trapper.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
10349636  
2  
3  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
11042159  
JOURNAL PUBMED  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, F., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)





Qy	841	AAAGGAATTTTGGGCGACCACTTTTGGATTACAGAGTCTCTTCTCTGACTATATCTCTCC	900
Db	2147	AAAGGAATTTTGGGCGACCACTTTTGGATTACAGAGTCTCTTCTCTGACTATACCCCTCC	2206
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATGTATGGATGCTCTCAAGCC	960
Db	2207	AGAAATTTTCTCTTTTGAAGTACTACTGATTTACATGTATGGATGCTCTCAAGCC	2266
Qy	961	TCATGATCTACAGCCCTGGAAAGAAATATCTTACTGTCTGTTCATATATATGGTCTCTCA	1020
Db	2267	TCATGATCTACAGCCCTGGAAAGAAATATACCCCACTGTATTATCATATATGTGTGCTCCA	2326
Qy	1021	-----	1020
Db	2327	GGTGACGTGGTGAAACAATCGGTTTAAAGGAGTCAAGTATTTCCGCGCTGAACACCCCTGGC	2386
Qy	1021	-----	1020
Db	2387	CTCCCTGGGTTATGTGGTTGTGGTATAGACACAGGGGATCTGTCTACCGAGGACTTAA	2446
Qy	1021	-----	1020
Db	2447	ATTGGAAGCGCCTTTAAATATAAAATGGTCAAAATAGAAATCGATCAAGTGAAGG	2506
Qy	1021	-----	1020
Db	2507	ACTCCAGTACTAGCATCTCAGTATGACTTTCATTGGACTTGGATCGAGTGGCATCCACGG	2566
Qy	1021	-----	1020
Db	2567	CTGGTCTTATGGTGGTACCTCTCCCTGATGGCATTAATGCAGAGGTCGGATCTTCGG	2626
Qy	1021	GGTGTCTATTTGTGGGCGCCAGTCACTCTGTGTGATCTTCTATGATACAGGATACAGGA	1080
Db	2627	GGTGTCTATTTGTGGGCGCCAGTCACTCTGTGTGATCTTCTATGATACAGGATACAGGA	2686
Qy	1081	AGTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTAAGGATCTGTGGCCAT	1140
Db	2687	GCCTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTAAGGATCTGTGGCCAT	2746
Qy	1141	GCAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTACTCTCTTACATGGTTTCTCTGA	1200
Db	2747	GCAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTACTCTCTTACATGGTTTCTCTGA	2806
Qy	1201	TGAGATGTCCTATTTGCACATACAGTATATTAAGTATTTTGTAGTGGGCTGGAAA	1260
Db	2807	TGAGATGTCCTATTTGCACATACAGTATATTAAGTATTTTGTAGTGGGCTGGAAA	2866
Qy	1261	GCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGATTAAGGATCTCTGATTCG	1320
Db	2867	GCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGATTAAGGATCTCTGATTCG	2926
Qy	1321	AGAACATTTATGAACCTGCACTTTTGCACCTACCTTCAAGAAACCTTTGGATCAGTATTCG	1380
Db	2927	AGAACATTTATGAACCTGCACTTTTGCACCTACCTTCAAGAAACCTTTGGATCAGTATTCG	2986
Qy	1381	TGCTCTTAAAGTATATTAATTTTCACTGTGTAGTACTCTCTGTGTATACATGGCTATT	1440
Db	2987	TGCTCTTAAAGTATATTAATTTTCACTGTGTAGTACTCTCTGTGTATACATGGCTATT	3039
Qy	1441	TAACCAATGAGGAGTTTATCAACAGAAACACAGATTTGATCATCAATTTGATACC	1500
Db	3040	TAACCAATGAGGAGTTTATCAACAGAAACACAGATTTGATCATCAATTTGATACC	3099
Qy	1501	TGCCATGTAAATCTCTCTGAAATAAATGTGGTGGCATGC	1543
Db	3100	TGCCATGTAAATCTCTCTGAAATAAATGTGGTGGCATGC	3142

RESULT 3  
AY411615  
LOCUS

DEFINITION	Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY411615		
VERSION	AY411615.1 GI:39767583		
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 2649) Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M. Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios		
TITLE	Science 302 (5652), 1960-1963 (2003)		
JOURNAL	14671302		
PUBMED	2 (bases 1 to 2649)		
REFERENCE	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M. Direct Submission Submitted (18-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..2649		
gene	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>2649 /gene="DPP8" /locus_tag="HCM4283"		
ORIGIN	Query Match 53.2%; Score 888.4; DB 10; Length 2649; Best Local Similarity 70.6%; Pred. NO. 1.5e-218; Matches 1199; Conservative 0; Mismatches 200; Indels 300; Gaps 1;		
Qy	1	AAACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGTCTCAGAAATATGATTCGATCTGA	60
Db	951	AAACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGTCTCAGAAATATGATTCGATCTGA	1010
Qy	61	AGGAAGGATCATAGATGTCTAGATAGGAACCTTCAACCTTTTGAGATTCCTATTGTA	120
Db	1011	AGGAAGGATCATAGATGTCTAGATAGGAACCTTCAACCTTTTGAGATTCCTATTGTA	1070
Qy	121	AGGAGTTGATATATTTGCCAGAGCTGGATGGACTCTCGAGGAAATATGCTTGGTCCAT	180
Db	1071	NN	1130
Qy	181	CCTACTAGATCGCTCCAGACTCCGCTACAGATAGTGTTCATCTACCTGAATTTATTTAT	240
Db	1131	CCTACTAGATCGCTCCAGACTCCGCTACAGATAGTGTTCATCTACCTGAATTTATTTAT	1190
Qy	241	CCGAGTGAAGATGATGTTATGGAAGGAGAGCTCATTGAGTCTGCTGATTCCTGT	300
Db	1191	CCGAGTGAAGATGATGTTATGGAAGGAGAGCTCATTGAGTCTGCTGATTCCTGT	1250
Qy	301	GAGGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT	360
Db	1251	GAGGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT	1310
Qy	361	TCATGTTTTCCTCAAGTACAGAGGAAATTTGAGTTTATTTTTCCTCTGAAATGCAA	420
Db	1311	TCATGTTTTCCTCAAGTACAGAGGAAATTTGAGTTTATTTTTCCTCTGAAATGCAA	1370
Qy	421	AAACAGGTTTCTGATTTATATCAAAATTTACATCTATTTTAAAGGAAGCAAAATATAACG	480



Db 1371 AACAGGTTCCGTCATTTATACAAATTAATCATCTATTTTAAAGGAAGCAATATAAAG 1430  
Qy 481 ATCCAGTGGTGGGCTGCTGCTCCAGAGTGAITTCGAAGTGTCTATCAAGAGGAGATAGC 540  
Db 1431 ATCCAGTGGTGGGCTGCTGCTCCAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1490  
Qy 541 AATTAACAGTGGTGAATGGGAAGTTCCTGGCCGGCATGGATCTAATATCCAGTTGATGA 600  
Db 1491 NN 1550  
Qy 601 AGTCAGAGGCTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGGTA 660  
Db 1551 AGTCAGAGGCTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGGTA 1610  
Qy 661 CGTAGTCAGTTACGTAATCTGGAGAGGTGA CAAGGCTGACTGACCGTGCTACTCACA 720  
Db 1611 CGTAGTCAGTTACGTAATCTGGAGAGGTGA CAAGGCTGACTGACCGTGCTACTCACA 1670  
Qy 721 TTCTTGCTGATCAGTCAGCATGTGCACTTCTTTATAAGTAAAGTATAGTAAACCAAGAA 780  
Db 1671 TTCTTGCTGATCAGTCAGCATGTGCACTTCTTTATAAGTAAAGTATAGTAAACCAAGAA 1730  
Qy 781 TCACACCTGTGTCTCCCTTTACAGCTATCAAGTCTGAAGTACCCCACTTCGCAAAAC 840  
Db 1731 TCACACCTGTGTCTCCCTTTACAGCTATCAAGTCTGAAGTACCCCACTTCGCAAAAC 1790  
Qy 841 AAAGGAATTTTGGGCCACCACTTTTGGATTGACAGGCTCTCTCTGACTATATCTCTCC 900  
Db 1791 AAAGGAATTTTGGGCCACCACTTTTGGATTGACAGGCTCTCTCTGACTATATCTCTCC 1850  
Qy 901 AGAAATTTCTTTTGAAGTACTACTGGAATTAATGATGATGATGATGATGATGATGATG 960  
Db 1851 AGAAATTTCTTTTGAAGTACTACTGGAATTAATGATGATGATGATGATGATGATGATG 1910  
Qy 961 TCATGATCTACAGCTGGAAGAAATATCTACTGCTGTGTTTCATATATGATGATGATGATG 1020  
Db 1911 TCATGATCTACAGCTGGAAGAAATATCTACTGCTGTGTTTCATATATGATGATGATGATG 1970  
Qy 1021 ----- 1020  
Db 1971 GGTGACGTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATACCTTAGC 2030  
Qy 1021 ----- 1020  
Db 2031 CTCTTAGGTTATGTTGTTAGTAGACACAGGGGATCTGTACCAGGGCTTAA 2090  
Qy 1021 ----- 1020  
Db 2091 ATTTGAAGGCGCTTTAAATATATAAATGGTCAAAATGGAATTCAGCATCAGGTGGAGG 2150  
Qy 1021 ----- 1020  
Db 2151 ACTCAATATCTAGTTCTCGATATGATGATTTTCAATTCAGTTAGATCGTGGGCATCCAGG 2210  
Qy 1021 ----- 1020  
Db 2211 CTGTCCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTCAG 2270  
Qy 1021 GGTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATCAGATACAGGA 1080  
Db 2271 GGTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATCAGATACAGGA 2330  
Qy 1081 ACGTTATATGGTCAACCTGACAGCAATCAACAGGCTATTAATCTAGGATCTGTGGGCAT 1140  
Db 2331 ACGTTATATGGTCAACCTGACAGCAATCAACAGGCTATTAATCTAGGATCTGTGGGCAT 2390  
Qy 1141 GCAAGCAGAAAAGTTCCCTCTGAACCAATCTGTTTACTGCTCTTACATGTTTCTCGGA 1200  
Db 2391 GCAAGCAGAAAAGTTCCCTCTGAACCAATCTGTTTACTGCTCTTACATGTTTCTCGGA 2450  
Qy 1201 TGAGATGTCATTTTGCACATACAGATATATCTAGTATTTTTTGTAGGGCTGAAA 1260  
Db 2451 TGAGATGTCATTTTGCACATACAGATATATCTAGTATTTTTTGTAGGGCTGAAA 2510

Qy 1261 GCCATATGATTTACAGATCTATCTCCTCAGGAGAGACACAGCATAGAGTTCTGATCGG 1320  
Db 2511 GCCATATGATTTACAGATCTATCTCCTCAGGAGAGACACAGCATAGAGTTCTGATCTGG 2570  
Qy 1321 AGACATATTAACATGCTATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGCTATTGC 1380  
Db 2571 AGACATATTAACATGCTATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGCTATTGC 2630  
Qy 1381 TGCTCTAAAAGTGATATAA 1399  
Db 2631 TGCTCTAAAAGTGATATAA 2649  
RESULT 4  
AK029788  
LOCUS  
DEFINITION  
Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930560C15 product:DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens], full insert sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AK029788.1 GI:26081520  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
10349636  
2  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
3 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
11042159  
4  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
5 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
11076861  
6  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
7 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
8  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
9 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
10  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
11 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, K., Saiton, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,  
Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayaishizaki, Y.

# TITLE JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayaishizaki, The Institute of  
Physical and Chemical Research (RIKEN) Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,  
Kanagawa 230-0045, Japan. (E-mail: genome-res@gscc.riken.jp,  
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

URL: <http://fantom.gsc.riken.jp/>

## FEATURES Source

1. 5517  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:4930560C15"  
/db\_xref="taxon:10090"  
/clone="4930560C15"  
/sex="male"  
/tissue\_type="testis"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
2221\_4837  
misc\_feature  
[note="DIPEPTIDYL PEPTIDASE 8 homolog (Homo sapiens)  
(SPTR109HBM5, evidence: PASTY, 95.7%ID, 100%length,  
match=2649)  
putative"]

## ORIGIN

Query Match 52.7%; Score 879.4; DB 4; Length 5517;  
Best Local Similarity 72.6%; Pred. No. 3.8e-216;  
Matches 1362; Conservative 0; Mismatches 201; Indels 312; Gaps 5;  
QY 67 GATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATTCTATTGAAGAGT 126  
DB 3205 GATTATAGATGTCATAGATAAGAACTGGTTCACCTTTTCGAGATTCTGTTGAGGAGT 3264  
QY 127 TGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATATGCTTGGTCCATCTTACT 186  
DB 3265 TGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATATGCTTGGTCCATCTTACT 3324  
QY 187 AGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTATTATCCAGT 246  
DB 3325 AGACCGTTCCAGACTCACCTGCAGATAGTTCGTATCTCCCTCGAGTTATTATCCAGT 3384  
QY 247 AGAAGATGATGTTATGGAAGGCAGAGACTCAATTGAGTCAGTGCCTGATCTGTGAGCC 306  
DB 3385 AGAAGATGATGCCATGGAAGGCAGAGACTTATAGAGTCAGTTCCTGACTCTGTGACACC 3444  
QY 307 ACTAATTATCTAGGAACACAGACATCTGGTAATATCCATGACATCTTCTCATGT 366  
DB 3445 ACTGATCATCTATGAGAAACACAGACATCTGGTAATATCCACGATATTTTTCATGT 3504  
QY 367 TTTTCCCAAGTCAGAGAGAAATGAGTTATTTTTCCTCTGAGTGCAAAACAGG 426  
DB 3505 TTTTCTCAAACTCATGAAGATGAATTTGAGTTATTTTTCCTCTGAGTGCAAAACAGG 3564  
QY 427 TTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAAAATATAAGATCCAG 486  
DB 3565 TTTTCGTCTATCTGTATAAATACATCTATTTTAAAGGAGAGCAAAATATAAGATCCAG 3624  
QY 487 TGTGGGCTGCTGCTCCAGTGATTTTCAAGTGTCTTATCAAGAGGAGATAGCAATTAC 546  
DB 3625 TGTGGACTACCTGCCCCAAGTGATTTCAAGTGTCTTATCAAGGAAGAAATAACAATTAC 3684

QY 547 CAGTGTGAATGGGAAGTTCCTTGGCCGCGCATCGATCTAATATCCAAAGTTGATGAGTCA 606  
DB 3685 CAGTGTGAATGGGAAGTTCCTTGGCCGCGCATCGATCTAATATCTGGTTGATGAAGCCAG 3744  
QY 607 AAGCTGGTATATTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGT 666  
DB 3745 AAGCTGGTATATTGAAGGACCAAGACTTCCTTTGGAGCATCACCTGTAGCTAGC 3804  
QY 667 CAGTTACGTAAATCTCGAGAGGTGACAAAGGTGACTGACCGTGGTCTACTACATCTCT 726  
DB 3805 CAGTTATGCAAAACCTCGAGAGGTGAGGAGTGTGAGGCTGACTGACCGTGGTCTACT 3864  
QY 727 CTGCATCAGTCAGACATCTGACTCTTTTATAAGTAAGTATAGTAAACCAAGAAATCC 786  
DB 3865 CTGCCTCAGCCGGCATTTGACTCTTCTATAAGTAAGTACAGCAACCAAGAAATCC 3924  
QY 787 CTGTGTCTCCTTTTCAAAAGCTATCAAGTCTCAAGATGACCCCAACTTGCACAAAC 846  
DB 3925 CTGTGTCTCCTCTCAAAACTCTCAAGTCTCAGGATGACCCAGTTCATAAACAAAG 3984  
QY 847 ATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTCTGACTATATCTCTCC 906  
DB 3985 ATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTCTGACTACACCCCTCC 4044  
QY 907 TTTCTCTTTTGAAGTACTACTGGATTTTACATTTGTATGGATGCTCTACAAGCCCTCA 966  
DB 4045 TTTTCTTTTGAAGTACTACTGGATTTTACATCTGTATGGAATGTTGTATAGCCCTCA 4104  
QY 967 TCTACAGCTCGGAAGAAATATCTTACTGCTGTGTTTCATATATGCTGCTCAGGTT-- 1024  
DB 4105 CTTAAACCTTGAAGAAATATACCCCACTGTGTATTTCATATATGTTGTTGCCAGGTG 4164  
QY 1025 ----- 1024  
DB 4165 CTGCTGNAACAATCGCGTTTAAAGGAGTCAAGTATTTCCGCTGGAACACCCCTGGC 4224  
QY 1025 ----- 1024  
DB 4225 GGGTTATGTGGTTGTGGTATAGACAAAGGGGATCTCTGTCAACGAGGACTTAAATTT 4284  
QY 1025 ----- 1024  
DB 4285 AGCGCCTTTAAATATAAAATGGGTCAAAATAGAAATCGATGATCAAGTGGAGGACT 4344  
QY 1025 ----- 1024  
DB 4345 GTACTAGCATCTCAGTATGACTTCAATGACTTGGATCGAGTGGGCATCCACGGCTG 4404  
QY 1025 -----GC 1026  
DB 4405 CTATGGTGGCTACCTCTCCTGATGGCAATTAATGCAGAGGTCCGATATCTTCCGGT 4464  
QY 1027 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTT 1086  
DB 4465 TATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAGC 4524  
QY 1087 TATGGTTCACCTGACAGATGAAACAGGCTATTACTTAGGATCTGTGGCATGCAAGC 1146  
DB 4525 TATGGTTCACCTGACAGATGAAACAGGCTATTACTTAGGATCTGTGGCATGCAAGC 4584  
QY 1147 AGAAAGTTCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAA 1206  
DB 4585 GGAAGTTCCTCTGAAACCAAAACCGGTTACTCTCTTGTATGGTTCTTGGATGAA 4644  
QY 1207 TGTCCATTTTGCATACATCAGTATATTAATCTAGATTTTGTAGGGCTGGAAGGCCATA 1266  
DB 4645 TGTTCATTTTGCACACACAGTATATGCTGAGTTTGTAGGGCTGGAAGGCCATA 4704  
QY 1267 TGAATTTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCTGTAATCGGAGAAC 1326  
DB 4705 TGACTTACAGATCTATCTCTCAGGAGGACACAGCATGAGTTCTGAGTCTGTGAGAAC 4764

QY 1327 TTATGAAGTGCATCTTTTGGACCTTACCTTCAAGAAACCTTGGATCACGATTTGCTGCTCT 1386  
|||||  
Db  
4765 CTATGAATGCACCTGCTCCACCTTCCAGGAGAACTTGGATCGCGTATCGCTGCTCT 4824  
|||||  
QY 1387 AAAAGTGATATAATTTTGGACCTGTGTAGAACTCTCTGGGTATACACTGGCTATTAAACAA 1446  
|||||  
Db 4825 GAAAGTGATATACTCTGACGATGTAGAAGTCTCTGGGTATACACTGGCTATTAAACAA 4877  
|||||  
QY 1447 ATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCATCTTTTGTATACCTGGCAT 1506  
|||||  
Db 4878 ACAAGAGGGTTAGGCAATAGAGAGCCAGAGTTGATGAACCTACTTTTGTATACCTGGCAT 4937  
|||||  
QY 1507 GTAACATCTACTCTGAAATTAATGTGTGTCATGCGAGGGGTCTACGGTTTGGG-TAG 1565  
|||||  
Db 4938 ATAAGATCTATTTCTGAAAGTAAATTTGGTGCATGCGAGAAATCTGCGAGTTTCATGSGATAG 4997  
|||||  
QY 1566 TAATCTAATACCTTAACCCACATGTCACAAATCAAAATGATACATATTCCTCGAGAGACCC 1625  
|||||  
Db 4998 TACTCTAATACCTTAC- - - -CATGCATATAAAGTGTGATGATGTTCCCAAGGATTTG 5053  
|||||  
QY 1626 AGCAATACCATAGA 1640  
|||||  
Db 5054 AACAGTACCTGAGGA 5068  
|||||

RESULT 5  
AY411616 2649 bp DNA linear GSS 16-DEC-2003  
LOCUS Pan troglodytes Dpp8 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY411616  
VERSION AY411616.1 GI:39767584  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Pan.  
REFERENCE 1 (bases 1 to 2649)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 2649)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source 1..2649  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>2649  
/gene="Dpp8"  
/locus\_tag="HCM4283"  
ORIGIN  
Query Match 50.5%; Score 842.8; DB 10; Length 2649;  
Best Local Similarity 68.1%; Pred. No. 9.6e-207;  
Matches 1157; Conservative 0; Mismatches 242; Indels 300; Gaps 1;  
QY 1 AACAGGTACAGCAAAATCCCTTAAGTCACTTTTAAGATGTGCAAGATAATGATGATGCTGA 60  
|||||

951 AACAGGTACAGCAAAATCCCTTAAGTCACTTTTAAGATGTGCAAGATAATGATGATGCTGA 1010  
|||||  
QY 61 AGSAGAGATCATAGATGCTCATAGATAAGGAATAATTTCAACCTTTTGAGATTTCTATTTGA 120  
|||||  
1011 AGSAGAGNN 1070  
|||||  
QY 121 AGSAGTTGAATATATTCGCCAGAGCTGGAGTCTCTGAGGGAATAATATGCTTGGTCCAT 180  
|||||  
1071 NNN 1130  
|||||  
QY 181 CCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTGTGTGATCTCACCTGGAATTTATTTAT 240  
|||||  
1131 CCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTGTGTGATCTCACCTGGAATTTATTTAT 1190  
|||||  
QY 241 CCCAGTGAAGATGATGTTATGAAAGGAGAGATCATTTAGTGTGAGTGTGCTGATTTCTGT 300  
|||||  
1191 CCCAGTGAAGATGATGTTATGAAAGGAGAGATCATTTAGTGTGAGTGTGCTGATTTCTGT 1250  
|||||  
QY 301 GACGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360  
|||||  
1251 GACGCCACTAATTTATCNCATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1310  
|||||  
QY 361 TCATGTTTTTCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGGCTCTGGAATGCAA 420  
|||||  
1311 TCATGTTTTTCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGGCTCTGGAATGCAA 1370  
|||||  
QY 421 AACAGGTTTCGTCATTTATACAAATATCATCTATTTTAAAGGAAAGCAATATAAAGC 480  
|||||  
1371 AACAGGTTTCGTCATTTATACAAATATCATCTATTTTAAAGGAAAGCAATATAAAGC 1430  
|||||  
QY 481 ATCCAGTGTGGGCTGCTGCTCCAGAGTATTTCAAGTGTCTCTATCAAGAGAGATAGC 540  
|||||  
1431 ATCCAGTGTGGGCTGCTGCTCCCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1490  
|||||  
QY 541 AATTACCAGTGTGAATGGGAAGTTCTTGGCCGCATGATCTAATATATCCAAAGTTGATGA 600  
|||||  
1491 NNN 1550  
|||||  
QY 601 AGTCAGAGGCTGTATATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCCTGTA 660  
|||||  
1551 AGTCAGAGGCTGTATATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCCTGTA 1610  
|||||  
QY 661 CGTAGTCAGTTACGTAATCTCTGAGAGGTGCAAGGCTGACTGACCGTGGTACTCACA 720  
|||||  
1611 CGTAGTCAGTTACGTAATCTCTGAGAGGTGCAAGGTTGACTGACCGTGGTACTCACA 1670  
|||||  
QY 721 TTCTTGTGTCATCAGTCAGCATGTCACCTTCTTTTAAAGTAAGTATAGTAACCAAGAA 780  
|||||  
1671 TTCTTGTGTCATCAGTCAGCATGTCACCTTCTTTTAAAGTAAGTATAGTAACCAAGAA 1730  
|||||  
QY 781 TCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAAC 840  
|||||  
1731 TCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAAC 1790  
|||||  
QY 841 AAAGGAATTTTGGGCGACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATACCTCTCC 900  
|||||  
1791 AAAGGAATTTTGGGCGACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATACCTCTCC 1850  
|||||  
QY 901 AGAAATTTCTCTTTTGAAGTACTACTGATTTACATGTTATGGATGCTCTACAAGCC 960  
|||||  
1851 AGAAATTTCTCTTTTGAAGTACTACTGATTTACATGTTATGGATGCTCTACAAGCC 1910  
|||||  
QY 961 TCATGATCTACAGCTCGAAAGAAATATCCTACTGTGCTGTTTATATATGTTGGTGGTCTCA 1020  
|||||  
1911 TCATGATCTACAGCTCGAAAGAAATATCCTACTGTGCTGTTTATATATGTTGGTGGTCTCA 1970  
|||||  
QY 1021 ----- 1020  
1971 GGTGAGTTGGTGAATAATCGGTTTAAAGGCAAGTATTTCCGCTTGAATACCTAGC 2030  
|||||  
QY 1021 ----- 1020  
2031 CTCTCTAGGTTATGTTGTTAGTGTATAGAACACAGGGGATCCTGTACCGAGGGCTTAA 2090  
|||||

QY	1021	-----	1020
Db	2091	ATTGTAAGGCGCTTTAAATATAAAATGGGTCAAATAGAAATTGACGATCAGGTGGAAGG	2150
QY	1021	-----	1020
Db	2151	ACTCCAATATCTAGCTTCTCGATATGATTTCATTGACTTAGATCGTGTGGGCATCCACGG	2210
QY	1021	-----	1020
Db	2211	CTGGTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTGAGATATCTTCAG	2270
QY	1021	GGTTGCTATTGCTGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	1080
Db	2271	GGTTGCTATTGCTGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	2330
QY	1081	ACGTTATATGGGTCAACCCTGACCAGAAATGAACAGGCGCTATTACTTAGGATCTGTGGCCAT	1140
Db	2331	ACGTTATATGGGTCAACCCTGACCAGAAATGAACANNNNNNNNNNNNNNNNNNNNNNNNN	2390
QY	1141	GCAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGA	1200
Db	2391	NNNNNNNNNNNNNNNNNNNNNNNACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGA	2450
QY	1201	TGAGATGCTCAATTTGCACATACCACTATATTAATGAGTTTCTGCTCTTACATGGTTTCTCTGGA	1260
Db	2451	TGAGATGCTCAATTTGCACATACCACTATATTAATGAGTTTCTGCTCTTACATGGTTTCTCTGGA	2510
QY	1261	GCCATATGATTTACAGATCTATCTCTGAGGAGACACAGCATAGAGTTCCTGAATCGGG	1320
Db	2511	ACCATATGATTTACAGATCTATCTCTGAGGAGACACAGCATAGAGTTCCTGAATCTGG	2570
QY	1321	AGAACTATTGAACTGATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGCATATGC	1380
Db	2571	AGAACTATTGAACTGATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGCATATGC	2630
QY	1381	TGCTCTAAAAGTGATATAA	1399
Db	2631	TGCTCTAAAAGTGATATAA	2649
RESULT 6			
BX372276/c			
LOCUS	BX372276 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED		
DEFINITION	Homo sapiens cDNA clone CS0DL005YD02 3-PRIME, mRNA sequence.		
ACCESSION	BX372276		
VERSION	BX372276.1 GI:30448117		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
	1. (bases 1 to 910)		
	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.		
	Full-length cDNA libraries and normalization		
	Unpublished (2001)		
COMMENT	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE		
	Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr		
	1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7542.r		
For more information about this cluster, see			
http://www.genoscope.cns.fr/cdna?s=CS0BAI043ZH06_CS04080_1&c=7542.r			
. FEATURES			
Location/Qualifiers			

source	1. 910									
	/organism="Homo sapiens"									
	/mol_type="mRNA"									
	/db_xref="taxon:9606"									
	/clone="CS0DL005YD02"									
	/cell_type="B CELLS (RAMOS CELL LINE)"									
	/cell_line="RAMOS CELL LINE"									
	/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"									
	/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."									
ORIGIN										
Query Match	49.2%;	Score 820.6;	DB 5;	Length 910;						
Best Local Similarity	99.0%;	Pred. No. 4.2e-201;								
Matches 856;	Conservative 0;	Mismatches 6;	Indels 3;	Gaps 3;						
QY	1	AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGA	60							
Db	870	AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGA	811							
QY	61	AGGAAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCATTATTTGA	120							
Db	810	AGGAAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCATTATTTGA	753							
QY	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCAGGAGGAAAATATGCTTGGTCCAT	180							
Db	752	AGNAGTTGGATATATTGCCAGAGCTGGATGGACTCTCAGGAGGAAAATATGCTTGGTCCAT	693							
QY	181	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTTGTGATCTCACCTGAAATTTATTTAT	240							
Db	692	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTTGTGATCTCACCTGAAATTTATTTAT	633							
QY	241	CCCAGTAAAGATGATGTTTATGAAAGGAGAGACTCAATTGAGTCAGTCGCTGATTCGT	300							
Db	632	CNCAGTAAAGATGATGTTTATGAAAGGAGAGACTCAATTGAGTCAGTCGCTGATTCGT	573							
QY	301	GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	360							
Db	572	GACGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT	513							
QY	361	TCATGTTTTTCCCAAAAGTCACGAGAGGAAATTCAGTTTATTTTTGGCTCTCGAATGCAA	420							
Db	512	TCATGTTTTTCCCAAAAGTCACGAGAGGAAATTCAGTTTATTTTTGGCTCTCGAATGCAA	453							
QY	421	AACAGGTTTCCGTCATTTATACAAATTAATACATCTATTTTAAAGGAAAGCAAAATATAACG	480							
Db	452	AACAGGTTTCCGTCATTTATACAAATTAATACATCTATTTTAAAGGAAAGCAAAATATAACG	393							
QY	481	ATCCAGTGGTGGGCTCGCTGCTCCAGTCAATTTCAAGTGTCTATCAAGAGGAGATAGC	540							
Db	392	ATCCAGTGGTGGGCTCGCTGCTCCAGTCAATTTCAAGTGTCTATCAAGAGGAGATAGC	333							
QY	541	AATTACAGTGGTGAATGGGAGTTCTTGGCCGGATGGATCTAATATCCAAAGTTGATGA	600							
Db	332	AATTACAGTGGTGAATGGGAGTTCTTGGCCGGATGGATCTAATATCCAAAGTTGATGA	273							
QY	601	AGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACTGTA	660							
Db	272	AGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACTGTA	213							
QY	661	CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGAACAAGGCTGACTGACCGTGGCTACTACA	720							
Db	212	CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGAACAAGGCTGACTGACCGTGGCTACTACA	153							
QY	721	TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAAACGAAGAA	780							
Db	152	TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAAACGAAGAA	93							
QY	781	TCCACACTGTGTGTCCCTTTTACAAGCTATCAAGTCTCGAAGATGACCCAACTTGGAAAAAC	840							

```

Db      92  TCCACACTGTGTCCCTTTACAAGCTATCAAGTCTGAAGATGACCCAAAC-TGCCAAAC 34
Qy      841 AAAGGAATTTTGGGCCACCAATTTG 865
Db      33  AAAGGAATTTTGGGCCACCAATTTG 9

RESULT 7
DN517062/c
LOCUS      DN517062      902 bp      mRNA      linear      EST 10-MAR-2005
DEFINITION 1257572 MARC 7BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION  DN517062
VERSION     DN517062.1  GI:60727252
KEYWORDS
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1 (bases 1 to 902)
AUTHORS     Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
            Wray,J.E. and Keefe,J.W.
TITLE       A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL
COMMENT     Unpublished (2003)
            Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: RLK8033 row: G column: 12
            Seq primer: TAGAGGCGACAGTCGAGG.

FEATURES             source
            1..902      Location/Qualifiers
                        /organism="Bos taurus"
                        /mol_type="mRNA"
                        /db_xref="taxon:9913"
                        /tissue_type="pooled"
                        /lab_host="DH10B"
                        /clone_lib="MARC 7BOV"
                        /notes="vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
                        Library made with RNA pooled from multiple tissues
                        including ovary, hindbrain, uterus, and day-30 whole
                        embryos."

ORIGIN
Query Match      48.2%; Score 804.4; DB 8; Length 902;
Best Local Similarity 93.2%; Pred. No. 6.e-197;
Matches 841; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy      4  AGGTACAGCAAACTCTTAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAAGG 63
Db      902 AGGCACAGCAAACTCAAAAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAAGG 843

Qy      64  AAGGATCATAGATGTCATAGATAAGCAACTAATTCACCTTTTGAGATTCCTATTGTAAGG 123
Db      842 AAGGATTTATAGATGTCATAGATAAGCAACTAATTCAGCCTTTTGAGATTCCTATTGTAAGG 793

Qy      124 AGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAATATGCTTGGTCCATCCT 183
Db      782 AGTTGAATATATTGCCAGAGCTGGATGGACTCCAGGGGAAAATATGCTGCGGCATCCT 723

Qy      184 ACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATATTATTATCCC 243
Db      722 ACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATATTATTATCCC 663

Qy      244 AGTAGAGATGATGTTATGTGAAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTCGTGAC 303
Db      662 AGTAGAGATGATGTCATGTGAAAGACAGAGATTTCATCGAGTCAGTGCCTGATTCGTGAC 603

```

```

Qy      304 GCCACTAATATTCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGATCTTTCA 363
Db      502 ACCACTGATTAATCTATGAAGAAACGACAGACATCTGGATAAATATCCATGATCTTTCA 543

Qy      364 TGTTCCTCCCAAGTCAAGAGAGAAATTGAGTTTATTTTGGCTCTGAATCAAAAC 423
Db      542 TGTTCCTCCCAAGTCAAGAGATGAATTTGAATTTATTTTGGCTCTGAATCAAAAC 483

Qy      424 AGTTTCCTGCTATTTATACAAAATACATCTATTTTAAAGGAAAGCAATATAACCATC 483
Db      482 AGTTTCCTGCTATTTATATAAATACATCCATTTTAAAGGAGAGCAAGTATAAGCGATC 423

Qy      484 CAGTGTGGGCTGCTGCTCCCAAGTCAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAAT 543
Db      422 CAGTGTGGGCTGCTGCTGCTCCCAAGTCAATTTCAAGTGTCTCTATTAAGAGGAAATAGCAAT 363

Qy      544 TACCAGTGTGTGAATGGAAAGTCTTGGCGGGCATCGATCTAATATCCAAAGTTGATGAAGT 603
Db      362 TACCAGTGTGTGAATGGAAAGTCTTGGCGGGCATCGATCTAATATCCAAAGTTGATGAAGT 303

Qy      604 CAGAAGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTAGCT 663
Db      302 CAGAAGCTGGTATATTTTGAAGGCAACCAAGACTCTCTTTAGAACATCACCTGTATGT 243

Qy      664 AGTCAGTTAGCTAAATCCTGGAGAGGTGACAGGCTGACCTGCGTGGCTACTCACATTC 723
Db      242 GGTCAAGTTATGTAATCTTGGAGAGGTGACAGGCTGACCTGCGTGGTATTTCGCACTC 183

Qy      724 TTGCTGCATCAGTCAGCACTGTGACTTCTTTAATAGTAAGTATAGTACCAAGAAATCC 783
Db      182 TTGCTGCATCAGCGCGCATTTGACTTCTTTAATAGTAAGTATAGTACCAAGAGAGGCC 123

Qy      784 ACACGTGTCTCCCTTTTACAAGCTATCAAGTCTCTCAAGATGACCCAACTTCGCAAAACAA 843
Db      122 ACACGTGTCTCCCTTTTACAAGCTCTCAAGTCTCTCAAGCAGCACCACCTGCAAGACAA 63

Qy      844 GGAATTTTGGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTGATCATATCTCTCCAGA 903
Db      62 AGAATTCCTGGCCACCGCTCTTGGATTCAGCAGGTCTCTTCTGATCATATACCCCTCCAAA 3

Qy      904 AA 905
Db      2 AA 1

RESULT 8
DN519771
LOCUS      DN519771      857 bp      mRNA      linear      EST 11-MAR-2005
DEFINITION 1262448 MARC 7BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION  DN519771
VERSION     DN519771.1  GI:60958943
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1 (bases 1 to 857)
AUTHORS     Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
            Wray,J.E. and Keefe,J.W.
TITLE       A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL
COMMENT     Unpublished (2003)
            Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: RLK8033 row: G column: 12
            Seq primer: GTAATACGACTCACTATAGG.

```

```
FEATURES
  source
    Location/Qualifiers
      1..857
        /organism="Bos taurus"
        /mol_type="mRNA"
        /db_xref="taxon:9913"
        /tissue_type="pooled"
        /lab_host="DH10B"
        /clone_lib="MARC 780V"
        /notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
        library made with RNA pooled from multiple tissues
        including ovary, hindbrain, uterus, and day-30 whole
        embryos."
      ORIGIN
        Query Match 43.8%; Score 730.6; DB 8; Length 857;
        Best Local Similarity 93.9%; Pred. No. 8.5e-178;
        Matches 760; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
        Qy 1 AACAGGTCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTCAGATCTCTATTGGA 60
        Db 49 AACAGGCACAGCAAAATCCAAAAGTCACCTTTTAAGATGTCAGAAAATCTGATGACGCTGA 108
        Qy 61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTCAGATCTCTATTGGA 120
        Db 109 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTCAGATCTCTATTGGA 168
        Qy 121 AGGAGTTGAATATATTCAGAGAGCTGGATGGAATCTCTGAGGGAATAATGCTTGGCCAT 180
        Db 169 AGGAGTTGAATATATTCAGAGAGCTGGATGGAATCTCTGAGGGAATAATGCTTGGCCAT 228
        Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTAT 240
        Db 229 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTAT 288
        Qy 241 CCCAGTAGAGATGATGTTATCGGAAGCGAGACTCATTTGAGTCAGTGCCTGATCTGT 300
        Db 289 CCCAGTAGAGATGATGCTCATGGAAGACAGAGATTCATCGAGTCAGTGCCTGATCTGT 348
        Qy 301 GAGCCACTAATTAATCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360
        Db 349 GACACCACTGATTAATCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTT 408
        Qy 361 TCATGTTTTTCCCAAGTCACAGAAAGGAAATTTGAGTTTATTTTTGCTCTGAAATGCAA 420
        Db 409 TCATGTTTTTCCCAAGTCATGAAGATGAAATTTGAAATTTATTTTTGCTCTGAAATGCAA 468
        Qy 421 AACAGTTTCCGTCAATTTATACAAATTTACATCTATTTTAAAGGAAACGAAATATAAAGC 480
        Db 469 AACAGTTTCCGTCAATTTATATAAATACATCTCAATTTTAAAGGAGCAAGTATAAGCG 528
        Qy 481 ATCCAGTGTGGGCTGCCTGCTCCAAAGTGATTTTCAAGTGTCTATCAAAGAGGAGATAGC 540
        Db 529 ATCCAGTGTGGGCTGCCTGCTCCAAAGTGATTTTCAAGTGTCTATCAAAGAGGAAATAGC 588
        Qy 541 AATTACAGTGTGAATGGAAAGTTCTTTGGCCGCATGGATCTAATATCCAGTTGATCA 600
        Db 589 AATTACAGTGTGAATGGAAAGTTCTTTGGCCGCATGGATCTAATATCCAGTTGATCA 648
        Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTCTGA 660
        Db 649 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCTCTTTAGACATCACTCTGA 708
        Qy 661 CGTAGTCAGTACGTAATCTCTGAGAGGTGACAGAGCTGACACGCGGTCACTCA 720
        Db 709 TGTGTCAGTATGTAATCTCTGAGAGGTGACAGAGCTGACACGCGGTATTCGA 768
        Qy 721 TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTAACAGAGAA 780
        Db 769 CTCCTGTGTCATCAGCCGGCATTTGTGACTTCTTTAAGTAAGTATAGTAACAGAGAG 828
        Qy 781 TCCACACTGTGTGTCCTTTACAGCTAT 809
        Db 829 CCCACACTGTGTGTCCTTTACAGCTGT 857
```

```
RESULT 9
B1223892
LOCUS
DEFINITION
  B1223892 Homo sapiens CDNA clone IMAGE:5104113 5',
  mRNA sequence.
ACCESSION
  B1223892
VERSION
  B1223892.1 GI:14677336
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominiidae; Homo.
  1 (bases 1 to 853)
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: ATCC
  CDNA Library Preparation: Life Technologies, Inc.
  CDNA Library Arrayed by: Incyte Genomics, Inc.
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LHAM11251 row: e column: 10
  High quality sequence stop: 718.
  Location/Qualifiers
    1..853
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5104113"
      /tissue_type="cervical carcinoma cell line"
      /lab_host="DH10B"
      /clone_lib="NIH MGC_12"
      /notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
      Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
      Average insert size 1.4 kb. Library prepared by Life
      Technologies."
```

```
FEATURES
  source
    Location/Qualifiers
      1..853
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5104113"
        /tissue_type="cervical carcinoma cell line"
        /lab_host="DH10B"
        /clone_lib="NIH MGC_12"
        /notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
        Average insert size 1.4 kb. Library prepared by Life
        Technologies."
```

```
ORIGIN
  Query Match 41.2%; Score 68814; DB 2; Length 853;
  Best Local Similarity 99.7%; Pred. No. 7.3e-167;
  Matches 700; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
  Qy 323 GAAACACACAGACATCTGGATAAATATCCATGACATCTTTCA-TGTTTTTCCCAAGTCA 381
  Db 1 GAAACACACAGACATCTGGATAAATATCCATGACATCTTTCAAGTGTTTTCCCAAGTCA 60
  Qy 382 CGAAGAGGAAATTTGAGTTTATTTTTCCTCTGAATGCAACAGGTTTCGGTCATTTATA 441
  Db 61 CGAAGAGGAAATTTGAGTTTATTTTTCCTCTGAATGCAACAGGTTTCGGTCATTTATA 120
  Qy 442 CAAATTTACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGTGGCTGCTGC 501
  Db 121 CAAATTTACATCTATTTTAAAGGAAGCAATATAACGATCCAGTGGTGGCTGCTGC 180
  Qy 502 TCCAGTGAATTTCAAGTGTCTTATCAAAGAGAGATAGCAATTAACAGTGGTGAATGGGA 561
  Db 181 TCCAGTGAATTTCAAGTGTCTTATCAAAGAGAGATAGCAATTAACAGTGGTGAATGGGA 240
  Qy 562 AGTTCTTGGCCGGCATGGAATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTT 621
  Db 241 AGTTCTTGGCCGGCATGGAATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTT 300
  Qy 622 TGAAGCACCACAAAGACTCCCTTTTAGAGCATCACCTGTAGTCAGTACGTAAATCC 681
  Db 301 TGAAGCACCACAAAGACTCCCTTTTAGAGCATCACCTGTAGTCAGTACGTAAATCC 360
```



Qy 682 TGGAGAGTGACAAAGCTGACCTGACGCGTGGCTACTCAATCTTGTGCTGATCAGTACGCA 741  
Db |||||  
Qy 361 TGGAGAGTGACAAAGCTGACCTGACGCGTGGCTACTCAATCTTGTGCTGATCAGTACGCA 420  
Db |||||  
Qy 742 CTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTA 801  
Db |||||  
Qy 421 CTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTA 480  
Db |||||  
Qy 802 CAAGCTATCAAGTCTCTGAAGATGACCCAACTTTCGCAAAAACAAAGGAATTTTGGGCGCCAT 861  
Db |||||  
Qy 481 CAGCTATCAAGTCTCTGAAGATGACCCAACTTTCGCAAAAACAAAGGAATTTTGGGCGCCAT 540  
Db |||||  
Qy 862 TTTGGATTACGAGGCTCTCTTCGACTATATCTCTCCAGAAATTTCTTTTGAAG 921  
Db |||||  
Qy 541 TGTGGATTACGAGGCTCTCTTCGACTATATCTCTCCAGAAATTTCTTTTGAAG 600  
Db |||||  
Qy 922 TACTACTGGATTACATTTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAA 981  
Db |||||  
Qy 601 TACTACTGGATTACATTTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAA 660  
Db |||||  
Qy 982 GAAATATCCTACTGTGCTGTTTCATATATGTTGCTCTCAGGT 1023  
Db |||||  
Qy 661 GAAATATCCTACTGTGCTGTTTCATATATGTTGCTCTCAGGT 702  
Db |||||

RESULT 10  
AV411617 2634 bp DNA linear GSS 16-DEC-2003  
LOCUS Mus musculus DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION AV411617  
VERSION AV411617.1 GI:39767585  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridea; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2634)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adam,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 2634)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adam,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment  
FEATURES  
source  
1. .2634  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
gene  
1. .2634  
/gene="DPP8"  
/locus\_tag="HCM4283"

ORIGIN  
Query Match 40.5%; Score 675.6; DB 10; Length 2634;  
Best Local Similarity 72.1%; Pred. No. 2e-163;  
Matches 738; Conservative 0; Mismatches 285; Indels 0; Gaps 0;  
Qy 1 AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATTGATGCTGA 60  
|||||



```

LOCUS      CN647232      1013 bp      mRNA      linear      EST 13-MAY-2004
DEFINITION ILLUMIGEN MCQ 28042 Katze MMR Macaca mulatta cDNA clone IBIUW:7816
            5' similar to bases 1 to 1009 highly similar to human DPP8
            (Hs.439202), mRNA sequence.
VERSION    CN647232
KEYWORDS   CN647232.1 GI:47160675
SOURCE     EST
ORGANISM   Macaca mulatta (rhesus monkey)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE  1 (bases 1 to 1013)
AUTHORS   Magnes,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proli,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
TITLE      Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
JOURNAL    Genome Biol. 6 (7), R60 (2005)
PUBMED     15998449
COMMENT    Contact: C. Magnes
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagnes@illumigen.com
            PCR Primers
            Sequenced on 2004.03.20. 793 Q20 bases.
            FORWARD: CCTCACTAAAGGACAA
            BACKWARD: CACTATAGGCGAATTGGTA
            Insert Length: 1013 Std Error: 0.00
            Plate: CL000179 Row: E Column: 08
            Seq primer: CCTCACTAAAGGACAA
POLYA=No. Location/Qualifiers
            1..1013
            /organism="Macaca mulatta"
            /mol_type="mRNA"
            /strain="Indian"
            /db_xref="taxon:9544"
            /clone="IBIUW:7816"
            /sex="female"
            /dev_stage="adult"
            /lab_host="E. coli SOLR"
            /notes="Organ: brain; Vector: Uni-ZAP XR; Site: 1: EcoR I;
            Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis
            kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold
            Cloning Kit (Catalog #200450)"

ORIGIN
Query Match      40.3%; Score 672.6; DB 7; Length 1013;
Best Local Similarity 98.7%; Pred. No. 9.4e-163;
Matches 678; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      337 CTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCACGAAGGAAATTGA 396
DB      1 CTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCACGAAGGAAATTGA 60
QY      397 GTTATTTTGTCTCGAATCAAAACAGGTTTCGGTCATTATACAAATTTACATCTAT 456
DB      61 GTTATTTTGTCTCGAATCAAAACAGGTTTCGGTCATTATACAAATTTACATCTAT 120
QY      457 TTATAGGAAGCAATATAACGATCCAGTGGTGGCTGCTCCCAAGTCAATTCAA 516
DB      121 TTATAGGAAGCAATATAACGATCCAGTGGTGGCTGCTCCCAAGTCAATTCAA 180
QY      517 GTGTCTATCAAGAGGAGATAGCAATTTACAGTGGTGGATGGGAAGTTCTTGGCCGGCA 576
DB      181 GTGTCTATCAAGAGGAGATAGCAATTTACAGTGGTGGATGGGAAGTTCTTGGCCGGCA 240
QY      577 TGGATCTAATATCCAGTTGTAAGTGAAGTCAAGGCTGGTATATTTTGAAGGCACCAAGA 636
            |||

```

```

DB      241 TGGATCTAATATCCAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 300
QY      637 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTTAAATCTCTGGAGAGGTGACAAG 696
DB      301 CTCCTCTTTAGAGCATCATCTGTACGTAGTCAGTTACGTTAAATCTCTGGAGAGGTGACAAG 360
QY      697 GCTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAGCAGCTGTGACTTCTTTAT 756
DB      361 GCTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAGCAGCTGTGACTTCTTTAT 420
QY      757 AAGTAAGTATAGTAACACAGAAAGTAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 816
DB      421 AAGTAAGTATAGTAACACAGAAAGTAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTTC 480
QY      817 TGAAGATGACCCAACTTGCACAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 876
DB      481 TGAAGATGATCTAACTTGCACAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 540
QY      877 TCCTCTTCTGACTATACCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTCAC 936
DB      541 TCCTCTTCTGACTATACCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTCAC 600
QY      937 ATTCTATGGGATCTCTTACAAAGCTCTATCATGATCTACAGCTGGAGAAATATCTCTACTGT 996
DB      601 ATTCTATGGGATCTCTTACAAAGCTCTATCATGATCTACAGCTGGAGAAATATCTCTACTGT 660
QY      997 GCTGTTTCATATATGGTGGTCTCTCAGGT 1023
DB      661 GCTGTTTCATATATGGTGGTCTCTCAGGT 687

RESULT 12
LOCUS      CN646612
DEFINITION ILLUMIGEN MCQ 26659 Katze MMR Macaca mulatta cDNA clone IBIUW:8421
            5' similar to bases 1 to 1025 highly similar to human DPP8
            (Hs.439202), mRNA sequence.
VERSION    CN646612
KEYWORDS   CN646612.1 GI:47160055
SOURCE     EST
ORGANISM   Macaca mulatta (rhesus monkey)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE  1 (bases 1 to 1027)
AUTHORS   Magnes,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
            Proli,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
            Iadonato,S.P.
TITLE      Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
JOURNAL    Genome Biol. 6 (7), R60 (2005)
PUBMED     15998449
COMMENT    Contact: C. Magnes
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagnes@illumigen.com
            Sequenced on 2004.03.10. 784 Q20 bases.
            PCR Primers
            FORWARD: CCTCACTAAAGGACAA
            Insert Length: 1027 Std Error: 0.00
            Plate: CL000146 Row: D Column: 07
            Seq primer: CCTCACTAAAGGACAA
POLYA=No. Location/Qualifiers
            1..1027
            /organism="Macaca mulatta"
            /mol_type="mRNA"
            /strain="Indian"
            /db_xref="taxon:9544"

```

```
/clone="IBIUM:8421"
/sex="female"
/dev_stage="adult"
/lab_host="E. coli SOLR"
/clone_lib="Katze MBR"
/notes="Organ: brain; Vector: Uni-ZAP XR; Site 1: Ecor I;
Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis
kit (Catalog #200400) and ZAP-cDNA Gigapack III Gold
Cloning Kit (Catalog #200450)"

ORIGIN
Query Match 40.3%; Score 672.6; DB 7; Length 1027;
Best Local Similarity 98.7%; Pred. No. 9.4e-163;
Matches 678; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 337 CTGGATAAATATCCATGACATCTTTCATGTTTTTCCCAAGTCCAGAGAGGAATTTGA 396
DB 1 CTGGATAAATATCCATGACATCTTTCATGTTTTTCCCAAGTCCAGAGAGGAATTTGA 60
QY 397 GTTATTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCATTTATACAAAATTACATCTAT 456
DB 61 GTTATTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCATTTATACAAAATTACATCTAT 120
QY 457 TTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGCTGCTCTCAAGTGATTTCAA 516
DB 121 TTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGCTGCTCTCAAGTGATTTCAA 180
QY 517 GTGTCTTATCAAGAGAGATAGCAATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCA 576
DB 181 GTGTCTTATCAAGAGAGATAGCAATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCA 240
QY 577 TGGATCTAATATCCAGTGTAGTACAGTGGTGTATATTTGAAGCCACCAAGA 636
DB 241 TGGATCTAATATCCAGTGTAGTACAGTGGTGTATATTTGAAGCCACCAAGA 300
QY 637 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTTACGTTAAATCTCGAGAGGTGACAAG 696
DB 301 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTTACGTTAAATCTCGAGAGGTGACAAG 360
QY 697 GCTGACTGACCGTGGCTACTCACAATTTTGTGCAATCAGTCCAGTGGTACTTCTTTAT 756
DB 361 GCTGACTGACCGTGGCTACTCACAATTTTGTGCAATCAGTCCAGTGGTACTTCTTTAT 420
QY 757 AAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 816
DB 421 AAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTTC 480
QY 817 TGAAGATGCCAACCTTGCAAAACAGGAATTTTGGGCCACCAATTTTGGATTCCAGCAGG 876
DB 481 TGAAGATGATCTAACTTGCAAAACAGGAATTTTGGGCCACCAATTTTGGATTCCAGCAGG 540
QY 877 TCCTCTCTCTGACTATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTAC 936
DB 541 TCCTCTCTCTGACTATACCCCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTAC 600
QY 937 ATTGTATGGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAGAATAATCTTACTGT 996
DB 601 ATTGTATGGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAGAATAATCTTACTGT 660
QY 997 GCTGTTTCATATATGGTGGTCTCTCAGGT 1023
DB 661 GCTGTTTCATATATGGTGGTCTCTCAGGT 687

RESULT 13
BG259714
LOCUS 993 bp mRNA linear EST 13-FEB-2001
DEFINITION 602380072F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4511080 5',
mRNA sequence.
ACCESSION BG259714
VERSION BG259714.1 GI:12769530
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
1 (bases 1 to 993)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10393 row: 0 column: 17  
High quality sequence start: 6  
High quality sequence stop: 656.  
Location/Qualifiers  
1. .993  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4511080"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_92"  
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dr primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 39.5%; Score 660; DB 2; Length 993;  
Best Local Similarity 100.0%; Pred. No. 1.7e-159;  
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 277 CATTGAGTCAGTCGCTGATTCGTGACGCCACTAATTATCTATGAAGAAACACAGACAT 336  
DB 1 CATTGAGTCAGTCGCTGATTCGTGACGCCACTAATTATCTATGAAGAAACACAGACAT 60  
QY 337 CTGATATAATCCATGACATCTTTCATGTTTTTCCCAAGTCCAGAGAGGAATTTGA 396  
DB 61 CTGATATAATCCATGACATCTTTCATGTTTTTCCCAAGTCCAGAGAGGAATTTGA 120  
QY 397 GTTATTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCAATTTATACAAAATTACATCTAT 456  
DB 121 GTTATTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCAATTTATACAAAATTACATCTAT 180  
QY 457 TTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGCTGCTCTCAAGTGATTTCAA 516  
DB 181 TTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGCTGCTCTCAAGTGATTTCAA 240  
QY 517 GTGTCTTATCAAGAGAGATAGCAATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCA 576  
DB 241 GTGTCTTATCAAGAGAGATAGCAATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCA 300  
QY 577 TGGATCTAATATCCAGTGTAGTACAGTGGTGTATATTTGAAGCCACCAAGA 636  
DB 301 TGGATCTAATATCCAGTGTAGTACAGTGGTGTATATTTGAAGCCACCAAGA 360  
QY 637 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTTACGTTAAATCTCGAGAGGTGACAAG 696  
DB 361 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTTACGTTAAATCTCGAGAGGTGACAAG 420  
QY 697 GCTGACTGACCGTGGCTACTCACAATTTTGTGCAATCAGTCCAGTGGTACTTCTTTAT 756  
DB 421 GCTGACTGACCGTGGCTACTCACAATTTTGTGCAATCAGTCCAGTGGTACTTCTTTAT 480  
QY 757 AAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 816

```

Db      481  AAGTAAGTAGTAGTAACCAAGAGAAATCCACACTGTGTGTCCTTTTACAGACTATCAAGTCC 540
Qy      817  TGAAGATGACCACTTGCACCAAGGAAATTTTGGCCACCAATTTGGATTCAGCAGG 876
Db      541  TGAAGATGACCACTTGCACCAAGGAAATTTTGGCCACCAATTTGGATTCAGCAGG 600
Qy      877  TCCTCTTCTCGACTATACCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTAC 936
Db      601  TCCTCTTCTCGACTATACCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTAC 660

RESULT 14
AF175225 1292 bp mRNA linear HTC 01-AUG-2003
LOCUS Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.
AF175225
ACCESSION AF175225.1 GI:33338055
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1292)
Zhao,B., Xu,H.S., Tong,Y.K., Sheng,H., Qin,B.M., Liu,Y.Q., Liu,B.,
Wang,X.Y., Zhang,Q., Song,L., Gao,Y., Zhang,C.L., Ye,J., Ji,X.J.,
Liu,B.H., Lu,H., Chen,J.Z., Cai,M.Q., Zheng,W.Y., Teng,C.Y.,
Liu,Q., Yu,L.T., Lin,J., Gong,Q., Zhang,A.M., Gao,R.L. and Hui,R.T.
Direct Submission
Submitted (04-AUG-1999) Molecular Medicine Center for
Cardiovascular Disease, Cardiovascular Institute, CIMS & PUMC, 167,
Bei Li Shi Lu, Beijing 100037, P.R. China
FEATURES
Source Location/Qualifiers
1..1292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
142..1023
/codon_start=1
/product="MSTP135"
/protein_id="AA013650.1"
/db_xref="GI:33338056"
/translation="MTQLAKRNFGPPFWIQVLFLLILQKFLSKLVLLDLHMGCS
TSPHSTGLERILLXLFYIGPKVQLVNNRFGKFRFLRLTLFFLGVVXGIDNRGS
CRGLKFAKFGQIEIDQVEGLQYLASRYDFIDLRVGIHGSYGGYLSLMAL
MRSIDIFRVAIGAVLWIFDYGTIERYHVPDQNEQGYLGSVAMQAEKFPSEPN
RLLHLGLDENVHFAVTLISLLSFLVRAGKPYDLQIYQERHSIRVPESGEHYELHLL
HILQENLGSRIAALKVI"
ORIGIN
Query Match 38.7%; Score 646.4; DB 4; Length 1292;
Best Local Similarity 99.8%; Pred. No. 6.1e-156;
Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1021 GGTGTCTATTGTGGGGCCCCAGTCACCTCTCTGGATCTTCTATGATACAGGATACACGGA 1080
Db 645 GGTGTCTATTGTGGGGCCCCAGTCACCTCTCTGGATCTTCTATGATACAGGATACACGGA 704
Qy 1081 AGCTTATGGGTGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
Db 705 AGCTTATGGGTGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 764
Qy 1141 GCAAGCAGAAAAGTTCCCTCTGCAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 1200
Db 765 GCAAGCAGAAAAGTTCCCTCTGCAACCAATCGTTTACTGCTCTTACATGGTTTCTCTGGA 824
Qy 1201 TGAGAAATGCCATTTTGGACATACAGTATATTACTAGTATTTTGTAGTGGGCTGGAA 1260
Db 825 TGAGAAATGCCATTTTGGACATACAGTATATTACTAGTATTTTGTAGTGGGCTGGAA 884
Qy 1261 GCCATATGATTTTACAGATCTATCTCTCGAGAGACACAGATAAGAGTTCCTGAATCGG 1320

```

```

Db      895  GCCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGAGTTCCTGAATCGGG 944
Qy 1321 AGAACATATTGAATCGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTAATGC 1380
Db 945 AGAACATATTGAATCGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTAATGC 1004
Qy 1381 TGCTCTAAAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440
Db 1005 TGCTCTAAAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1064
Qy 1441 AACCAAAATGAGGAGGTTTAAATCAACAGAAAAACACAGAATTGATCATCATTTTTCATACC 1500
Db 1065 AACCAAAATGAGGAGGTTTAAATCAACAGAAAAACACAGAATTGATCATCATTTTTCATACC 1124
Qy 1501 TGCCATGTAACTACTCTCTGAAAAATAAATGTGGTGCATCGACGGGTCTACGGTTTGT 1560
Db 1125 TGCCATGTAACTACTCTCTGAAAAATAAATGTGGTGCATCGACGGGTCTACGGTTTGT 1184
Qy 1561 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGA 1620
Db 1185 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTTCTGAGA 1244
Qy 1621 GACCCAGCAATACCATTAAGAAATTTACTTAAAAAATAAATAAATAAATAAATAAATAA 1668
Db 1245 GACCCAGCAATACCATTAAGAAATTTACTTAAAAAATAAATAAATAAATAAATAAATAA 1292

RESULT 15
BU631054 722 bp mRNA linear EST 23-SEP-2002
LOCUS UI-H-FEI-bdc-j-05-0-UI.s1 NCI-CGAP_FBI Homo sapiens cDNA clone
DEFINITION UI-H-FEI-bdc-j-05-0-UI 3', mRNA sequence.
BU631054
ACCESSION
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE Hominidae; Homo.
1 (bases 1 to 722)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES
Source Location/Qualifiers
1..722
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cloned="UI-H-FEI-bdc-j-05-0-UI"
/tissue_type="Cell lines"
/development="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI-CGAP_FBI"
/note="Organ: Chondrosarcoma; Vector: p7T3-Pac
(pharmacia) with a modified polylinker; Site_1: Boor I;
Site_2: Not I; NCI-CGAP_FBI is a normalized cDNA library
derived from a pool of mRNA obtained from 3 cell lines
from grade II chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA"

```

synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCTACGGAC. The cell lines were provided by Dr James Martin from the University of Iowa.  
TAG\_TISSUE=Human grade 2 chondrosarcoma cell line pool  
TAG\_LIB=UI-H-FBI  
TAG\_SEQ=CGCTACGGAC"

ORIGIN

Query Match		38.6%	Score 644;	DB 5;	Length 722;
Best Local Similarity		100.0%;	Pred. No. 2.2e-155;	Indels 0;	Gaps 0;
Matches 644;		Conservative 0;	Mismatches 0;		
Qy	1021	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	1080		
Db	644	GGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	585		
Qy	1081	ACGTTATATGGGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT	1140		
Db	584	ACGTTATATGGGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCAT	525		
Qy	1141	GCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCCTGGA	1200		
Db	524	GCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCCTGGA	465		
Qy	1201	TGAGAATGTCCATTTTGCACATACCAAGTATATTACTGAGTTTTTTAGTGGGGCTGAAA	1260		
Db	464	TGAGAATGTCCATTTTGCACATACCAAGTATATTACTGAGTTTTTTAGTGGGGCTGAAA	405		
Qy	1261	GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGATTAAGAGTTCCTGAATCGGG	1320		
Db	404	GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGATTAAGAGTTCCTGAATCGGG	345		
Qy	1321	AGAACATTATGAAGTCACTCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGCTATTGC	1380		
Db	344	AGAACATTATGAAGTCACTCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGCTATTGC	285		
Qy	1381	TGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	1440		
Db	284	TGCTCTAAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT	225		
Qy	1441	AACCAATGAGGAGGTTTAATCAACAGAAAACACAGANTTGATCATCATTTTGATACC	1500		
Db	224	AACCAATGAGGAGGTTTAATCAACAGAAAACACAGANTTGATCATCATTTTGATACC	165		
Qy	1501	TGCCATCTAACATCTACTCTCGAAATAATGTGGTGCCATGCGGGGTCTACGGTTTGT	1560		
Db	164	TGCCATCTAACATCTACTCTCGAAATAATGTGGTGCCATGCGGGGTCTACGGTTTGT	105		
Qy	1561	GGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA	1620		
Db	104	GGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA	45		
Qy	1621	GACCCAGCAATACCAATGAATTTACTTAAAAAATAAAAAA	1664		
Db	44	GACCCAGCAATACCAATGAATTTACTTAAAAAATAAAAAA	1		

Search completed: May 4, 2006, 07:56:08  
Job time : 6266.09 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2006, 01:01:26 ; Search time 528.298 Seconds  
(without alignments)  
5615.673 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 1669  
Sequence: 1 aacaggtagcagaataccta.....aaaaaaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/1 COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5 COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/H COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq:\*
- 7: /cgn2\_6/prodata/1/ina/PP COMB.seq:\*
- 8: /cgn2\_6/prodata/1/ina/RE COMB.seq:\*
- 9: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1669	100.0	1669	3	US-10-070-464-6
2	1648.4	98.8	4523	3	US-09-976-674-8
3	1501	89.9	4676	3	US-09-976-674-20
4	1349.6	80.9	4829	3	US-09-976-674-12
5	1347	80.7	3120	3	US-10-070-464-2
6	1343.6	80.5	4685	3	US-09-976-674-22
7	1337	80.1	2797	3	US-09-976-594-1103
8	1085.8	65.1	2671	3	US-09-976-674-2
9	881	52.8	1197	3	US-10-070-464-4
10	765.2	45.8	4309	3	US-09-976-674-14
11	493.4	29.6	823	3	US-09-280-116-171
12	434.2	26.0	1083	3	US-10-070-464-8
13	313	18.8	2617	3	US-09-976-674-4
14	313	18.8	4076	3	US-09-976-674-32
15	313	18.8	4159	3	US-09-976-674-30
16	313	18.8	4219	3	US-09-976-674-28
17	313	18.8	4302	3	US-09-976-674-24
18	297	17.8	308	3	US-09-621-976-1787
19	252.6	15.1	612	3	US-09-392-184-31
20	250.6	15.0	4037	3	US-09-976-674-40
21	250.6	15.0	4120	3	US-09-976-674-38
22	250.6	15.0	4180	3	US-09-976-674-36
23	250.6	15.0	4263	3	US-09-976-674-34
24	123.2	7.4	2411	3	US-09-976-674-26

25	89	5.3	89	3	US-09-513-999C-16433	Sequence 16433, A
26	74.8	4.5	936	3	US-09-270-767-12940	Sequence 12940, A
27	45.6	2.7	2924	3	US-09-023-655-1026	Sequence 1026, Ap
28	45.6	2.7	2924	6	PCT-US93-07923-1	Sequence 1, Appli
29	45.6	2.7	3407	3	US-10-002-593-5	Sequence 5, Appli
30	45.6	2.7	3407	3	US-09-949-016-275	Sequence 275, App
31	45.6	2.7	3407	3	US-10-423-714-5	Sequence 5, Appli
32	45.6	2.7	3419	3	US-09-949-016-4579	Sequence 4579, Ap
33	45.4	2.7	1023	3	US-09-270-767-13726	Sequence 13726, A
34	42.8	2.6	480	3	US-09-270-767-684	Sequence 684, App
35	42.8	2.6	480	3	US-09-270-767-15966	Sequence 15966, A
36	42.4	2.5	972	2	US-07-915-934-1	Sequence 1, Appli
37	42.4	2.5	972	2	US-08-325-743-1	Sequence 22, Appli
38	41.2	2.5	1141	3	US-09-806-708B-22	Sequence 11817, A
39	40.8	2.4	124700	3	US-09-949-016-11817	Sequence 14339, A
40	40.8	2.4	124701	3	US-09-949-016-15439	Sequence 22, Appli
41	40.6	2.4	1141	3	US-09-806-708B-22	Sequence 2813, Ap
42	40.2	2.4	832	3	US-09-621-976-2813	Sequence 13081, A
43	40	2.4	708	3	US-09-270-767-13081	Sequence 9, Appli
44	39.8	2.4	1405	3	US-09-244-111-9	Sequence 64, Appli
45	39.2	2.3	1474	3	US-08-821-994-64	

ALIGNMENTS

RESULT 1

US-10-070-464-6

; Sequence 6, Application US/10070464

; Patent No. 6881564

; GENERAL INFORMATION:

; APPLICANT: ABBOTT, Catherine Anne

; APPLICANT: GORRELL, Mark Douglas

; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES

; FILE REFERENCE: GH-007

; CURRENT APPLICATION NUMBER: US/10/070,464

; CURRENT FILING DATE: 2002-03-07

; PRIOR APPLICATION NUMBER: PCT/AU00/01085

; PRIOR FILING DATE: 2000-09-11

; PRIOR APPLICATION NUMBER: AU PQ5709

; PRIOR FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: AU PQ2762

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 1669

; TYPE: DNA

; ORGANISM: Homo Sapiens

US-10-070-464-6

Query Match	100.0%;	Score 1669;	DB 3;	Length 1669;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1669;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	AACAGGTACAGCAAACTCTTTAAAGTCACTTTAAAGTGTGTCAGAAATAATGATTGATCTGA	60
Db	1	AACAGGTACAGCAAACTCTTTAAAGTCACTTTAAAGTGTGTCAGAAATAATGATTGATCTGA	60
Qy	61	AGGAAGGATCATAGATGTCTAGATAAGCACTTAATTCACCTTTGAGATTCCTATTGA	120
Db	61	AGGAAGGATCATAGATGTCTAGATAAGCACTTAATTCACCTTTGAGATTCCTATTGA	120
Qy	121	AGGAGTTGATATATTGCCAGAGCTGATGCTCTCTGAGGAAATAATGCTTGCTCCAT	180
Db	121	AGGAGTTGATATATTGCCAGAGCTGATGCTCTCTGAGGAAATAATGCTTGCTCCAT	180
Qy	181	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT	240
Db	181	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT	240
Qy	241	CCCAGTGAAGATGATGTTTATGAAAGGAGAGACTTCATTGATGTCAGTCTGATCTGT	300
Db	241	CCCAGTGAAGATGATGTTTATGAAAGGAGAGACTTCATTGATGTCAGTCTGATCTGT	300

```
Db 241 CCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGACTCATTGAGTCAGTGCCTGATTCTGT 300
Qy 301 GAGCCCACTAATTTATCTATGAGAAACACACACATCTGGATAAATATCCATGACATCTT 360
Db 301 GAGCCCACTAATTTATCTATGAGAAACACACACATCTGGATAAATATCCATGACATCTT 360
Qy 361 TCATGTTTTTCCCAAGTCACGAGAGAAATTCAGTTTATTTTTCGCTCTGATGCAA 420
Db 361 TCATGTTTTTCCCAAGTCACGAGAGAAATTCAGTTTATTTTTCGCTCTGATGCAA 420
Qy 421 AACAGGTTTCGCTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACG 480
Db 421 AACAGGTTTCGCTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACG 480
Qy 481 ATCCAGTGTGGCTGCCCTGCCCAAGTGATTTTCAAGTGTCTCTATCAAAAGAGAGATAGC 540
Db 481 ATCCAGTGTGGCTGCCCTGCCCAAGTGATTTTCAAGTGTCTCTATCAAAAGAGAGATAGC 540
Qy 541 AATTACAGTGTGGAATGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 600
Db 541 AATTACAGTGTGGAATGGAAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 600
Qy 601 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTAGAGCATCACCTGTA 660
Db 601 AGTCAGAAGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTAGAGCATCACCTGTA 660
Qy 661 CGTAGTCAGTTACCTAAATCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 720
Db 661 CGTAGTCAGTTACCTAAATCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 720
Qy 721 TTCTCTGTCATCAGTCAGCAGCTGAGCTCTTTTAAAGTAAGTATAGTAACCAAGAA 780
Db 721 TTCTCTGTCATCAGTCAGCAGCTGAGCTCTTTTAAAGTAAGTATAGTAACCAAGAA 780
Qy 781 TCACACTGTGTGCTTACAGCTATCAAGTCTGAGAGTACGCCAACCCTTGCAAAAC 840
Db 781 TCACACTGTGTGCTTACAGCTATCAAGTCTGAGAGTACGCCAACCCTTGCAAAAC 840
Qy 841 AAGGAATTTTGGGCGACATTTGGATTACAGAGGTCTCTTCTCTGACTATATCTCTCC 900
Db 841 AAGGAATTTTGGGCGACATTTGGATTACAGAGGTCTCTTCTCTGACTATATCTCTCC 900
Qy 901 AGAAATTTTCTCTTTTGAAGTACTA CTGGATTTACATTTGATGGAGTCTCTACAGCC 960
Db 901 AGAAATTTTCTCTTTTGAAGTACTA CTGGATTTACATTTGATGGAGTCTCTACAGCC 960
Qy 961 TCATGATCTACAGCTGGAAAGAAATATCTCTACTGTCTGTTTATATATGGTGGTCTCA 1020
Db 961 TCATGATCTACAGCTGGAAAGAAATATCTCTACTGTCTGTTTATATATGGTGGTCTCA 1020
Qy 1021 GGTGCTATTTGTTGGGCCCCAGTCACCTCTCTGATCTTCTCTGACTATATCTCTCC 1080
Db 1021 GGTGCTATTTGTTGGGCCCCAGTCACCTCTCTGATCTTCTCTGACTATATCTCTCC 1080
Qy 1081 AGCTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCAT 1140
Db 1081 AGCTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGCCAT 1140
Qy 1141 GCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGA 1200
Db 1141 GCAAGCAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGA 1200
Qy 1201 TCAGATGTCCTATTTGACATACAGTATATTTACTGAGTTTTTTTATGAGGGCTGGAAA 1260
Db 1201 TCAGATGTCCTATTTGACATACAGTATATTTACTGAGTTTTTTTATGAGGGCTGGAAA 1260
Qy 1261 GCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTTGAATCGGG 1320
Db 1261 GCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTTGAATCGGG 1320
Qy 1321 AGAACTATTAACATCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGATTTGC 1380
Db 1321 AGAACTATTAACATCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGATTTGC 1380
```

## RESULT 2

US-09-976-674-8

; Sequence 8, Application US/09976674

; Patent No. 6844180

; GENERAL INFORMATION:

; APPLICANT: Qi Steve

; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669

; CURRENT APPLICATION NUMBER: US/09/976,674

; CURRENT FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: US 60/240,117

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 4523

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-976-674-8

Query Match 98.8%; Score 1648.4; DB 3; Length 4523;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

```
Qy 1 AACAGGTACAGCAAAATCCTAAAGTCACCTTTAAAGATGTCAGAAATATGATGATGCTGA 60
Db 1164 AACAGGTACAGCAAAATCCTAAAGTCACCTTTAAAGATGTCAGAAATATGATGATGCTGA 1223
Qy 61 AGGAAGGATCATAGATGTCATAGATTAAGAACTAAATTCACCTTTGAGATTCCTATTGA 120
Db 1224 AGGAAGGATCATAGATGTCATAGATTAAGAACTAAATTCACCTTTGAGATTCCTATTGA 1283
Qy 121 AGGAGTTGANTATATTTGCCAGAGCTGGATGACTCTCTGAGGAAATATATGCTTGGTCCAT 180
Db 1284 AGGAGTTGANTATATTTGCCAGAGCTGGATGACTCTCTGAGGAAATATATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCCAGACTCGCCTACAGATAGTGTGATCTCACCTCAATATTAT 240
Db 1344 CCTACTAGATCGCTCCCAGACTCGCCTACAGATAGTGTGATCTCACCTCAATATTAT 1403
Qy 241 CCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGACTCATTTGAGTGCCTGATTCGT 300
Db 1404 CCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGACTCATTTGAGTGCCTGATTCGT 1463
Qy 301 GAGGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGGCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 1523
```



QY 361 TCATGTTTTCCTCCCAAGTACGAGAGGAAATTTGAGTTTATTTTTTGCTCTGATGCAA 420  
Db 1524 TCATGTTTTCCTCCCAAGTACGAGAGGAAATTTGAGTTTATTTTTTGCTCTGATGCAA 1583  
QY 421 AACAGGTTTCCTGCTATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAAACG 480  
Db 1584 AACAGGTTTCCTGCTATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAAACG 1643  
QY 481 ATCCAGTGGTGGGCTGCTGCTCCCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGC 540  
Db 1644 ATCCAGTGGTGGGCTGCTGCTCCCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGC 1703  
QY 541 AATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATCAAGTTGATGA 600  
Db 1704 AATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGAATCTAATATCAAGTTGATGA 1763  
QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTA 660  
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTA 1823  
QY 661 CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACA 720  
Db 1824 CGTAGTCAGTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACA 1883  
QY 721 TTCTTGCTGTCATCAGTCAGCACTGTGCTTTTATTAAGTAACTAGTAAACAGAGAA 780  
Db 1884 TTCTTGCTGTCATCAGTCAGCACTGTGCTTTTATTAAGTAACTAGTAAACAGAGAA 1943  
QY 781 TCACACATGTTGTCCTTTTACAAGTATCAAGTCTCGAAGATGACCCAACTTGCAGAAC 840  
Db 1944 TCACACATGTTGTCCTTTTACAAGTATCAAGTCTCGAAGATGACCCAACTTGCAGAAC 2003  
QY 841 AAAGGAATTTGGCCACCATTTTGGATTGACAGGTCCTCTCTGACTATACCTCC 900  
Db 2004 AAAGGAATTTGGCCACCATTTTGGATTGACAGGTCCTCTCTGACTATACCTCC 2063  
QY 901 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGTATGGATGCTCTACAAGCC 960  
Db 2064 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGTATGGATGCTCTACAAGCC 2123  
QY 961 TCATGATCTACAGCTGGAAAGAAATCTCTACTGTCTGTCTATATATGTTGGTCTCA 1020  
Db 2124 TCATGATCTACAGCTGGAAAGAAATCTCTACTGTCTGTCTATATATGTTGGTCTCA 2179  
QY 1021 GGTGTGATTTGCTGGGGCCCAAGTCACTCTGTGATCTTTCTATGATACAGGATACAGGA 1080  
Db 2180 GGTGTGATTTGCTGGGGCCCAAGTCACTCTGTGATCTTTCTATGATACAGGATACAGGA 2239  
QY 1081 AGTTATATGGGTACCCCTGACCAAGTGAACAGGGCTATTTACTTAGGATCTGTGGCCAT 1140  
Db 2240 AGTTATATGGGTACCCCTGACCAAGTGAACAGGGCTATTTACTTAGGATCTGTGGCCAT 2299  
QY 1141 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200  
Db 2300 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 2359  
QY 1201 TGAGATGTCATTTTGGACATACCAAGTATATTTACTGAGTTTATTTAGTGGGCTGGAAA 1260  
Db 2360 TGAGATGTCATTTTGGACATACCAAGTATATTTACTGAGTTTATTTAGTGGGCTGGAAA 2419  
QY 1261 GCATATGATTTTACAGATCTATCTCAGGAGAGACACAGATAAGATTTCTGATCGGG 1320  
Db 2420 GCATATGATTTTACAGATCTATCTCAGGAGAGACACAGATAAGATTTCTGATCGGG 2479  
QY 1321 AGAATATATGAACTGATCTTTTGGACTATCTTCAAGAAACCTTGGATCAGTATTCG 1380  
Db 2480 AGAATATATGAACTGATCTTTTGGACTATCTTCAAGAAACCTTGGATCAGTATTCG 2539  
QY 1381 TGCTCTAAAAGTGTATATTTTGGCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440  
Db 2540 TGCTCTAAAAGTGTATATTTTGGCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2599  
QY 1441 AACCBAATGAGGAGTTTAAATCAAGAAAAACAGAAATTTGATCATCATTTTGTATACC 1500

Db 2600 AACCAATGAGGAGTTTAAATCAACAGAAACAGAAATTTGATCATCATTTTGTATACC 2659  
QY 1501 TGCCATGTAAACATCTACTCTCTGAAATAAATGTGGCCATGACGGGTCTACGGTTTGT 1560  
Db 2660 TGCCATGTAAACATCTACTCTCTGAAATAAATGTGGCCATGACGGGTCTACGGTTTGT 2719  
QY 1561 GGTAGTAACTTAATACCTTTAAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGAGA 1620  
Db 2720 GGTAGTAACTTAATACCTTTAAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGAGA 2779  
QY 1621 GACCCAGCAATACCATAGAAATTTACTAANAAAAAAGGAAAAA 1668  
Db 2780 GACCCAGCAATACCATAGAAATTTACTAANAAAAAAGGAAAAA 2827  
  
RESULT 3  
US-09-976-674-20  
; Sequence 20, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 4676  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-20  
  
Query Match 89.9%; Score 1501; DB 3; Length 4676;  
Best Local Similarity 91.5%; Pred. No. 0;  
Matches 1663; Conservative 0; Mismatches 5; Indels 149; Gaps 1;  
  
QY 1 AACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGCTGA 60  
Db 1164 AACAGGTACAGCAAAATCTCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGCTGA 1223  
QY 61 AGAAGGATCATAGATGTCATAGATAAGTAAGTAATTTCAACCTTTTGAGATTTCTATTGA 120  
Db 1224 AGAAGGATCATAGATGTCATAGATAAGTAAGTAATTTCAACCTTTTGAGATTTCTATTGA 1283  
QY 121 AGAGTTGATATATTTGCCAGAGCTGGAATGGAATCTCTGAGGAAATAATGCTGGTCCAT 180  
Db 1284 AGAGTTGATATATTTGCCAGAGCTGGAATGGAATCTCTGAGGAAATAATGCTGGTCCAT 1343  
QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTAT 240  
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTAT 1403  
QY 241 CCCAGTAGAAGATGATGTTATGSAAGGAGAGATCAATGAGTCAGTGCCTGATTTCTGT 300  
Db 1404 CCCAGTAGAAGATGATGTTATGSAAGGAGAGATCAATGAGTCAGTGCCTGATTTCTGT 1463  
QY 301 GACGCCACTAATTTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 360  
Db 1464 GACGCCACTAATTTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 1523  
QY 361 TCATGTTTTTCCCCAAAGTCAGAGAGGAAATTTGAGTTTATTTTTTGCTCTGATGCAA 420  
Db 1524 TCATGTTTTTCCCCAAAGTCAGAGAGGAAATTTGAGTTTATTTTTTGCTCTGATGCAA 1583  
QY 421 AACAGGTTTCCTGCTATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAAACG 480

Db 1584 AACAGGTTTCGGTCATTTATACAAATTAATACATCTATTTTAAAGGAAAGCAATATAACG 1643  
Qy 481 ATCCAGTGGTGGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAAAGAGGAGATAGC 540  
Db 1644 ATCCAGTGGTGGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAAAGAGGAGATAGC 1703  
Qy 541 AATTACAGTGGTGGTGAATGGGAAGTTCTTTGGCCGGGATGGATCTAATATCCAAGTTGATGA 600  
Db 1704 AATTACAGTGGTGGTGAATGGGAAGTTCTTTGGCCGGGATGGATCTAATATCCAAGTTGATGA 1763  
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGGATCACTCTGTA 660  
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGGATCACTCTGTA 1823  
Qy 661 CGTAGTCAGTTACCTAAATCTGGAGAGGTGACAGGCTGACCGCTGGCTACTCA 720  
Db 1824 CGTAGTCAGTTACCTAAATCTGGAGAGGTGACAGGCTGACCGCTGGCTACTCA 1883  
Qy 721 TTCTTCTGTCATCAGTCAGCACTGTGATCTTTTAAAGTAAATGATATGATAGTAAACGAAGA 780  
Db 1884 TTCTTCTGTCATCAGTCAGCACTGTGATCTTTTAAAGTAAATGATATGATAGTAAACGAAGA 1943  
Qy 781 TCCACACTGTGTCCCTTTACAGCTATCAAGCTCTGAGATGACCCAACTTGCAGAAC 840  
Db 1944 TCCACACTGTGTCCCTTTACAGCTATCAAGCTCTGAGATGACCCAACTTGCAGAAC 2003  
Qy 841 RAAGGAATTTGGGCAATTTGGATTGAGAGGTCTCTTCTGACTATATCTCTCC 900  
Db 2004 RAAGGAATTTGGGCAATTTGGATTGAGAGGTCTCTTCTGACTATATCTCTCC 2063  
Qy 901 AGNAATTTCTCTTTGAAAGTACTACGATTTACATTTGATGATGGATGCTCTACAGCC 960  
Db 2064 AGNAATTTCTCTTTGAAAGTACTACGATTTACATTTGATGATGGATGCTCTACAGCC 2123  
Qy 961 TCAATGATCAAGCTGGAAGAAATATCTTCTGCTCTTCTATATATGGTGGTCT --- 1016  
Db 2124 TCAATGATCAAGCTGGAAGAAATATCTTCTGCTCTTCTATATATGGTGGTGGT 2183  
Qy 1017 ----- 1016  
Db 2184 CAATAGAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTC 2243  
Qy 1017 ----- 1016  
Db 2244 ATTGACTTAGATCGTGTGGGATCCAGGCTGGTCTCTATGGAGGATACCTCTCCCTGATG 2303  
Qy 1017 -----CTCAGTTGCTATTGCTGGGCCCCAGTCACTCTG 1051  
Db 2304 GCATTAATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGCCCCAGTCACTCTG 2363  
Qy 1052 TGGATCTTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCTGACCCAGAGTCAA 1111  
Db 2364 TGGATCTTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCTGACCCAGAGTCAA 2423  
Qy 1112 CAGGCTATTACTTAGATCTGTGGCCATGCAAGCAAAAGTTCCCTCTGACCAAT 1171  
Db 2424 CAGGCTATTACTTAGATCTGTGGCCATGCAAGCAAAAGTTCCCTCTGACCAAT 2483  
Qy 1172 CGTTTACTGCTTTTACATGGTTTCTGGATGAGATGTCATTTTGGCATACACAGTATA 1231  
Db 2484 CGTTTACTGCTTTTACATGGTTTCTGGATGAGATGTCATTTTGGCATACACAGTATA 2543  
Qy 1232 TTACTGATTTTATGATGAGGCTGGAAGCCATATGATTTACAGATCTATCTCCAGAG 1291  
Db 2544 TTACTGATTTTATGATGAGGCTGGAAGCCATATGATTTACAGATCTATCTCCAGAG 2603  
Qy 1292 AGACACAGCATAGAGTTCTGATCGGGAGAACATTTATGAACTGCATCTTTTGCATAC 1351  
Db 2604 AGACACAGCATAGAGTTCTGATCGGGAGAACATTTATGAACTGCATCTTTTGCATAC 2663  
Qy 1352 CTTCAAGAAACCTTGGATCACTATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTG 1411  
Db 2664 CTTCAAGAAACCTTGGATCACTATTTGCTGCTCTAAAGTGATATAATTTTGACCTGTG 2723

RESULT 4

US-09-976-674-12 ; Sequence 12, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976.674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 4829  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-12

Query Match 80.9%; Score 1349.6; DB 3; Length 4829;  
Best Local Similarity 84.5%; Pred. No. 0;  
Matches 1684; Conservative 0; Mismatches 4; Indels 302; Gaps 1;

Qy 1 AACAGGTACAGCAAAATCCTAAAGTCACCTTTAAAGATGTCAGAAATATGATGATGCTGA 60  
Db 1164 AACAGGTACAGCAAAATCCTAAAGTCACCTTTAAAGATGTCAGAAATATGATGATGCTGA 1223  
Qy 61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGAGATTTCTATTGA 120  
Db 1224 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGAGATTTCTATTGA 1283  
Qy 121 AGGATTCGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAATAATATGCTTGGTCCAT 180  
Db 1284 AGGATTCGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAATAATATGCTTGGTCCAT 1343  
Qy 181 CCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACTCCCTGGAATTTAT 240  
Db 1344 CCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACTCCCTGGAATTTAT 1403  
Qy 241 CCCAGTGAAGATGATGTTATGAAAGGAGAGAGACTTGTGAGTCACTGCTGATCTCTGT 300  
Db 1404 CCCAGTGAAGATGATGTTATGAAAGGAGAGAGACTTGTGAGTCACTGCTGATCTCTGT 1463  
Qy 301 GACGCCACTAATATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360  
Db 1464 GACGCCACTAATATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523

Qy	361	TCATGTTTTTCCCAAGTCA	CGAAGAGGAATTGAGTTTATTTTGGCTCTGAATGCAA	420
Db	1524	TCATGTTTTTCCCAAGTCA	CGAAGAGGAATTGAGTTTATTTTGGCTCTGAATGCAA	1583
Qy	421	AACAGGTTTCGGTCATTTA	CAAAATTAATACATCTATTTTAAAGGAAGCAAAATATAACG	480
Db	1584	AACAGGTTTCGGTCATTTA	CAAAATTAATACATCTATTTTAAAGGAAGCAAAATATAACG	1643
Qy	481	ATCCAGTGGTGGCTGCCTG	CTCCAAAGTGATTTCAAGTGTCTTATCAAAAGGAGGATAGC	540
Db	1644	ATCCAGTGGTGGCTGCCTG	CTCCAAAGTGATTTCAAGTGTCTTATCAAAAGGAGGATAGC	1703
Qy	541	AATTACCAAGTGTGAATGG	GAAGTTCTTGGCCGGCATGGATCTAATATCAAAAGTGTATGA	600
Db	1704	AATTACCAAGTGTGAATGG	GAAGTTCTTGGCCGGCATGGATCTAATATCAAAAGTGTATGA	1763
Qy	601	AGTCAGAGGCTGGTATAT	TTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA	660
Db	1764	AGTCAGAGGCTGGTATAT	TTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTA	1823
Qy	661	CGTAGTCAGTTACGTAAT	CTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA	720
Db	1824	CGTAGTCAGTTACGTAAT	CTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACA	1883
Qy	721	TTCTTGCTGCATCAGTC	CAGCAGTGTGACTTTCTTTAATAGTAACTATAGTAAACGAGA	780
Db	1884	TTCTTGCTGCATCAGTC	CAGCAGTGTGACTTTCTTTAATAGTAACTATAGTAAACGAGA	1943
Qy	781	TCCACACTGTGTGCTCTT	TACAGCTATCAAGTCTCAAGATGACCCCACTTGCAAAAC	840
Db	1944	TCCACACTGTGTGCTCTT	TACAGCTATCAAGTCTCAAGATGACCCCACTTGCAAAAC	2003
Qy	841	AAAGGAATTTTGGGCCAC	CACTTTTGGATTGACAGGTCTCTCTGCTGACTATATCTCTCC	900
Db	2004	AAAGGAATTTTGGGCCAC	CACTTTTGGATTGACAGGTCTCTCTGCTGACTATATCTCTCC	2063
Qy	901	AGAAATTTTCTCTTTTGA	AGTACTACTGGATTACATTTGATGGGATGCTCTACAAGCC	960
Db	2064	AGAAATTTTCTCTTTTGA	AGTACTACTGGATTACATTTGATGGGATGCTCTACAAGCC	2123
Qy	961	TCATGATCTACAGCCTGG	AAAGAAATATCTCTGCTGTCTATATATGATGGTGTCTCA	1020
Db	2124	TCATGATCTACAGCCTGG	AAAGAAATATCTCTGCTGTCTATATATGATGGTGTCTCA	2183
Qy	1021	-----	-----	1020
Db	2184	CAGGTGCAGTTGGTGAAT	AATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTA	2243
Qy	1021	-----	-----	1020
Db	2244	GCCTCTTAGTTATGTGTT	GTGTAGTGATAGACACAGGGGATCCTGTCCAGGAGGCTT	2303
Qy	1021	-----	-----	1020
Db	2304	AAATTTGAAGGCCCTTTT	AAATATAAATGGGTCAATAGAAATGACGATCAGGTGGAA	2363
Qy	1021	-----	-----	1020
Db	2364	GGACTCCAATATCTAGCT	TCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCAC	2423
Qy	1021	-----	-----	1020
Db	2424	GGCTGGTCTATGGAGGAT	ACCTCTCCCTGATGGCATTAAATGACAGGTGAGATATCTTC	2483
Qy	1021	--GGTGTCTATTGCTGGG	CCCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACAG	1078
Db	2484	AGGGTGTCTATTGCTGGG	CCCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACAG	2543
Qy	1079	GAACGTTATATGGGTGAC	CCCTGACCAAGTAAAGAGGCTATTACTTAGGATCTGTGGCC	1138
Db	2544	GAACGTTATATGGGTGAC	CCCTGACCAAGTAAAGAGGCTATTACTTAGGATCTGTGGCC	2603
Qy	1139	ATGCAAGCAGAAAGTTCC	CTCTGAA CCAATCGTTTACTGCTCTTACATGTTTCCCTG	1198

Db	2604	ATGCAAGCAGAAAGTTCC	CTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCTG	2663
Qy	1199	GATCAGAAATGTCCATTT	TGACATACCAATGATATTTAGTGTAGGCTGGA	1258
Db	2664	GATCAGAAATGTCCATTT	TGACATACCAATGATATTTAGTGTAGGCTGGA	2723
Qy	1259	AAGCCATATGATTTACAG	ATCTATCTCAGGAGAGACAGCATAGAGTTCTCGAATCG	1318
Db	2724	AAGCCATATGATTTACAG	ATCTATCTCAGGAGAGACAGCATAGAGTTCTCGAATCG	2783
Qy	1319	GGAGAACATTATGAAC	TGCACTCTTTCACCAAAACCTTTGGATCAGGTAT	1378
Db	2784	GGAGAACATTATGAAC	TGCACTCTTTCACCAAAACCTTTGGATCAGGTAT	2843
Qy	1379	GCTGCTCTAAAAGTGA	TATAATTTTGACCTGTGTAGAACTCTCTGGTATACATGGCTAT	1438
Db	2844	GCTGCTCTAAAAGTGA	TATAATTTTGACCTGTGTAGAACTCTCTGGTATACATGGCTAT	2903
Qy	1439	TTAACCAAAATGAGGAG	GTTTTAAATCAACAGAAACACAGAAATTTGATCACAATTTGATA	1498
Db	2904	TTAACCAAAATGAGGAG	GTTTTAAATCAACAGAAACACAGAAATTTGATCACAATTTGATA	2963
Qy	1499	CCTGCCATGTAAACAT	CTTACTCTCTGAAATAAATGTGGCCATGACGGGTCTACGGTTT	1558
Db	2964	CCTGCCATGTAAACAT	CTTACTCTCTGAAATAAATGTGGCCATGACGGGTCTACGGTTT	3023
Qy	1559	GTGGTAGTAACTAAAT	CACTTAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGA	1618
Db	3024	GTGGTAGTAACTAAAT	CACTTAACCCCATGCTCAAAATCAAAATGATACATATTTCTCTGA	3083
Qy	1619	GAGACCCAGCAATACCA	TAAAGTAATTTACTAAAAAATAAAAAAAAAAAAAAAAAAAAA	1668
Db	3084	GAGACCCAGCAATACCA	TAAAGTAATTTACTAAAAAATAAAAAAAAAAAAAAAAAAAAA	3133

RESULT 5

US-10-070-464-2  
; Sequence 2, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 3120  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-070-464-2

Query Match	80.7%;	Score 1347;	DB 3;	Length 3120;
Best Local Similarity	84.7%;	Pred. No. 0;		
Matches 1657;	Conservative 0;	Mismatches 0;	Indels 300;	Gaps 1;
Qy	1	AACAGGTACAGCAAAATCC	TAAAGTCACTTTTAAGATGTGAGAAATTAATGATGCTGA	60
Db	1164	AACAGGTACAGCAAAATCC	TAAAGTCACTTTTAAGATGTGAGAAATTAATGATGCTGA	1223
Qy	61	AGGAAGATCATAGATGTG	TATAGTAAGGAATTAATTAACCTTTTGAGATCTATTGTA	120
Db	1224	AGGAAGATCATAGATGTG	TATAGTAAGGAATTAATTAACCTTTTGAGATCTATTGTA	1283

```
Qy 121 AGGATTGAATATATTCCAGAGCTGGATGGACTCTCGAGGAAATATGCTTGGTCCAT 180
Db 1284 AGAGTTGAATATATTCCAGAGCTGGATGGACTCTCGAGGAAATATGCTTGGTCCAT 1343
Qy 181 CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACTGAAATATTTAT 240
Db 1344 CCTACTAGATCGCTCCAGACTCGGCTACAGATAGTGTGATCTCACTGAAATATTTAT 1403
Qy 241 CCCAGTAGAAGATGATGTTTATGGAAAGGAGAGACTCAATGAGTCAGTGGCTGATCTGT 300
Db 1404 CCCAGTAGAAGATGATGTTTATGGAAAGGAGAGACTCAATGAGTCAGTGGCTGATCTGT 1463
Qy 301 GAGCCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 360
Db 1464 GAGCCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTT 1523
Qy 361 TCATGTTTTTCCCAAGTCAGAGAGAGGAAATGAGTTTATTTTTCCTCTGAATGCAA 420
Db 1524 TCATGTTTTTCCCAAGTCAGAGAGAGGAAATGAGTTTATTTTTCCTCTGAATGCAA 1583
Qy 421 AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAAGCAATATAAACG 480
Db 1584 AACAGGTTTCCGTCATTTATACAAATATACATCTATTTTAAAGGAAAGCAATATAAACG 1643
Qy 481 ATCCAGTGTGGGCTGCCCTGCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC 540
Db 1644 ATCCAGTGTGGGCTGCCCTGCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC 1703
Qy 541 AATTACAGTGTGAAATGGGAAGTTCTTGGCCGGCATGGATCTAAATATCCAGTTGATGA 600
Db 1704 AATTACAGTGTGAAATGGGAAGTTCTTGGCCGGCATGGATCTAAATATCCAGTTGATGA 1763
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTA 660
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTA 1823
Qy 661 CGTAGTCAGTTACGTAAATTCCTGGAGAGGTGACAAAGCTGACTGACCGTGCTACTCACA 720
Db 1824 CGTAGTCAGTTACGTAAATTCCTGGAGAGGTGACAAAGCTGACTGACCGTGCTACTCACA 1883
Qy 721 TCTTGTGTCAGTCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 780
Db 1884 TCTTGTGTCAGTCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 1943
Qy 781 TCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAC 840
Db 1944 TCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAC 2003
Qy 841 AAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGACTATACTCTCTC 900
Db 2004 AAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGACTATACTCTCTC 2063
Qy 901 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATGTTATGGGATGCTCTACAAGCC 960
Db 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATGTTATGGGATGCTCTACAAGCC 2123
Qy 961 TCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTCTCATATATGTTGCTCTCA 1020
Db 2124 TCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTCTCATATATGTTGCTCTCA 2183
Qy 1021 ----- 1020
Db 2184 GGTGCAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCCTAGC 2243
Qy 1021 ----- 1020
Db 2244 CTCTCTAGTATTATGTTGTTAGTAGATAGACAACAGGGGATTCCTGTACCGAGGGCTTAA 2303
Qy 1021 ----- 1020
Db 2304 ATTTGAAGGCGCTTTTAAATATAAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAAGG 2363
Qy 1021 ----- 1020

2364 ACTCCAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGG 2423
1021 ----- 1020
2424 CTGGTCCCTATGGAGGATACCTCTCCCTGATGGCATTAATGCGAGGGTCAGATATCTTCCAG 2483
1021 GGTTCCTATTTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGAGATACACGGA 1080
2484 GGTTCCTATTTGCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGAGATACACGGA 2543
1081 ACCTTATATGGGTCACTCCCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140
2544 ACCTTATATGGGTCACTCCCTGACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2603
1141 GCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGTCTTACATGCTTCTCTGGA 1200
2604 GCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGTCTTACATGCTTCTCTGGA 2663
1201 TGAGATGTCCATTTTGGCACATACCATGATATTTACTGAGTTTCTTGTAGGGGCTGGAAA 1260
2664 TGAGATGTCCATTTTGGCACATACCATGATATTTACTGAGTTTCTTGTAGGGGCTGGAAA 2723
1261 GCCATATGATTTACAGATCTATCCCTCAGGAGAGACACAGCATAGAGTTCCCTGAATCGGG 1320
2724 GCCATATGATTTACAGATCTATCCCTCAGGAGAGACACAGCATAGAGTTCCCTGAATCGGG 2783
1321 AGACATTATGATCTGCATCTTTTGCATCTACCTTCAAGAAAACTTGGATCACCTATTGTC 1380
2784 AGACATTATGATCTGCATCTTTTGCATCTACCTTCAAGAAAACTTGGATCACCTATTGTC 2843
1381 TGCTCTAAAAGTGATATATTTTGTACCTGTGTGAGAACTCTCTGGTATACATGCTGCTATTT 1440
2844 TGCTCTAAAAGTGATATATTTTGTACCTGTGTGAGAACTCTCTGGTATACATGCTGCTATTT 2903
1441 AACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCATCATTTTGTATACC 1500
2904 AACCAATGAGGAGGTTTAAATCAACAGAAACACAGAAATGATCATCATCATTTTGTATACC 2963
1501 TGCCATGTAACTACTACTCTCTGAAAAATAAATGTGTGTCCTGCGGATGTCACGGTTGT 1560
2964 TGCCATGTAACTACTACTCTCTGAAAAATAAATGTGTGTCCTGCGGATGTCACGGTTGT 3023
1561 GGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTCGAGA 1620
3024 GGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTCGAGA 3083
1621 GACCCAGCAATACCAAGAAATTTACTAAAAA 1657
3084 GACCCAGCAATACCAAGAAATTTACTAAAAA 3120

RESULT 6
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-976-674-22

Query Match 80.5%; Score 1343.6; DB 3; Length 4685;  
Best Local Similarity 87.0%; Pred. No. 0;  
Matches 1589; Conservative 0; Mismatches 79; Indels 158; Gaps 4;

QY 1 AACAGGTACAGCAATCCCTAAAGTCACCTTTTAAAGATGTCAGAAATTAATGATGATCTGA 60  
DB 1164 AACAGGTACAGCAATCCCTAAAGTCACCTTTTAAAGATGTCAGAAATTAATGATGATCTGA 1223

QY 61 AGGAGGATCATAGATGTCATAGATAGAGAACTTAATCAACCTTTTGAGATTCATTTGA 120  
DB 1224 AGGAGGATCATAGATGTCATAGATAGAGAACTTAATCAACCTTTTGAGATTCATTTGA 1283

QY 121 AGGAGTTGAATATATTGCGAGCTGGATGGACTCCTGAGGAAAATATGCTTGGTCCAT 180  
DB 1284 AGGAGTTGAATATATTGCGAGCTGGATGGACTCCTGAGGAAAATATGCTTGGTCCAT 1343

QY 181 CCTACTAGATCCCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTAT 240  
DB 1344 CCTACTAGATCCCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTAT 1403

QY 241 CCCAGTAGAAGATGATGTTATGGAAGCAGAGACTCATTTGAGTCAGTGCCTGATTCGT 300  
DB 1404 CCCAGTAGAAGATGATGTTATGGAAGCAGAGACTCATTTGAGTCAGTGCCTGATTCGT 1463

QY 301 GAGGCCACTAATATTCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 360  
DB 1464 GAGGCCACTAATATTCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 1523

QY 361 TCATGTTTTTCCCAAGTCCAGAGAGAAATTTGAGTTATTTTTGCTCTGAATGCAA 420  
DB 1524 TCATGTTTTTCCCAAGTCCAGAGAGAAATTTGAGTTATTTTTGCTCTGAATGCAA 1583

QY 421 AACAGTTTCCGTCTATTTATACAAATTAATCATCTATTTTAAAGAAAGCAATATTAACG 480  
DB 1584 AACAGTTTCCGTCTATTTATACAAATTTACATCTATTTTAAAGAAAGCAATATTAACG 1643

QY 481 ATCCAGTGTGGGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGC 540  
DB 1644 ATCCAGTGTGGGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGC 1703

QY 541 AATTACAGTGTGTAATGGAAAGTTCTTTGGCGGCATGGATCTAATATCCAAGTTGATGA 600  
DB 1704 AATTACAGTGTGTAATGGAAAGTTCTTTGGCGGCATGGATCTAATATCCAAGTTGATGA 1763

QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACATCCCTTTTAGAGATCACTCTGA 660  
DB 1764 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACATCCCTTTTAGAGATCACTCTGA 1823

QY 661 CGTAGTCAGTTACGTAATCTCGGAGAGGTGACAGGCTGACCTGACCGTGGCTACTCACA 720  
DB 1824 CGTAGTCAGTTACGTAATCTCGGAGAGGTGACAGGCTGACCTGACCGTGGCTACTCACA 1883

QY 721 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAGTAAGTATAGTAAACAGAGAA 780  
DB 1884 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAGTAAGTATAGTAAACAGAGAA 1943

QY 781 TCACACTGTGTCCCTTTTACAGCTATCAAGTCTGAAGATGACCCAACTTGGAAGAAC 840  
DB 1944 TCACACTGTGTCCCTTTTACAGCTATCAAGTCTGAAGATGACCCAACTTGGAAGAAC 2003

QY 841 AAAGGAATTTGGGCCACCAATTTGGATTCTAGTCCCTCAGGTGCAATTTGGTAATATCGG 2063  
DB 2004 AAAGGAATTTGGGCCACCAATTTGGATTCTAGTCCCTCAGGTGCAATTTGGTAATATCGG 2123

QY 897 CTCCA-----GAAATTTTCTTTTGAAGTACTACTGGAATTTTACATTTGATGGGATGC 950  
DB 2064 TTTAAAGAGTCAAGTATTTCCGCTTGAACTACCTAGCTCTCTAGGTTATGTTGTTA 2123

QY 951 TCTACAGGCTCATGATCTACAG----- 973  
DB 2124 GTGATAGACACAGGGGATCTCTGTCAACGAGGCTTAAATTTGAAGCGCTTTTAAATAT 2183

QY 974 ----- 973  
DB 2184 AAAATGGGTCAAATAGAAATTCACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGA 2243

QY 974 -----CCTGGAAAATAATATCTACTGTGCTGTTTCATATATATGTTGGTCTC 1019  
DB 2244 TATGATTTTATTGACTTAGATCGTGTGGCATCCACGGCTGGTCTATATGGAGGATACCTC 2303

QY 1020 -----AGTTTGTCTTATTTGCTGGGGCCCCA 1042  
DB 2304 TCCTGATGGCATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCA 2363

QY 1043 GTCACCTCTGTGGATCTTCTATGATACAGGATACAGGAACTTTATATGGTTCACCTTGAC 1102  
DB 2364 GTCACCTCTGTGGATCTTCTATGATACAGGATACAGGAACTTTATATGGTTCACCTTGAC 2423

QY 1103 CAGAATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCT 1162  
DB 2424 CAGAATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCT 2483

QY 1163 GAACCAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAAATGTCATTTTGCACAT 1222  
DB 2484 GAACCAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAAATGTCATTTTGCACAT 2543

QY 1223 ACCAGTATATTAATGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTAT 1282  
DB 2544 ACCAGTATATTAATGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTAT 2603

QY 1283 CCTCAGAGAGACACAGCATTAAGATTTCTGATCGGAGAGAACATTTATGAATCTGATCTT 1342  
DB 2604 CCTCAGAGAGACACAGCATTAAGATTTCTGATCGGAGAGAACATTTATGAATCTGATCTT 2663

QY 1343 TTGCATCTACCTTCAAGAAACCTTGATCAGCTATTGCTCTTAAAGTGTATATAATTT 1402  
DB 2664 TTGCATCTACCTTCAAGAAACCTTGATCAGCTATTGCTCTTAAAGTGTATATAATTT 2723

QY 1403 TGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATC 1462  
DB 2724 TGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATC 2783

QY 1463 AACAGAAAACACAGAAATGATCATCACATTTTGATACCTGCGCATGTAACATCTACTCTG 1522  
DB 2784 AACAGAAAACACAGAAATGATCATCACATTTTGATACCTGCGCATGTAACATCTACTCTG 2843

QY 1523 AAAATAAATGTGTGTCATGTCAGGCTCTACGTTTGTGGTAGTAATCTAATACTTAAAC 1582  
DB 2844 AAAATAAATGTGTGTCATGTCAGGCTCTACGTTTGTGGTAGTAATCTAATACTTAAAC 2903

QY 1583 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCAATAAAT 1642  
DB 2904 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCAATAAAT 2963

QY 1643 TACTAAAAA 1668  
DB 2964 TACTAAAAA 2989

## RESULT 7

US-09-976-594-1103  
; Sequence 1103, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,403  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program

```
; SEQ ID NO 1103
; LENGTH: 2797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1
US-09-976-594-1103

Query Match      80.1%; Score 1337; DB 3; Length 2797;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1661; Conservative 0; Mismatches 0; Indels 304; Gaps 2;

Qy      1 AACAGGTACAGCAAAATCCCTAAAGTCACCTTTAAAGATGTCAGAAAATAATGATGCTGA 60
Db      |
Qy      831 AACAGGTACAGCAAAATCCCTAAAGTCACCTTTAAAGATGTCAGAAAATAATGATGCTGA 890
Db      |
Qy      61 AGAAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGGAGATTCCTATTGGA 120
Db      |
Qy      891 AGAAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGGAGATTCCTATTGGA 950
Db      |
Qy      121 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAAATA---TGCTTGGT 176
Db      |
Qy      951 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAAATAAGTGTCTTGGT 1010
Db      |
Qy      177 CCATCCTACTAGATGCTCCAGACTCGCCTACAGATAGTGTGGATCTCACCTGAATATAT 236
Db      |
Qy      1011 CCATCCTACTAGATGCTCCAGACTCGCCTACAGATAGTGTGGATCTCACCTGAATATAT 1070
Db      |
Qy      237 TTATCCCACTAGAGATGATGTTATGGAAGGCGAGAGACTCCTAGTGTGAGTGTGAGTGTGAGT 296
Db      |
Qy      1071 TTATCCCACTAGAGATGATGTTATGGAAGGCGAGAGACTCCTAGTGTGAGTGTGAGTGTGAGT 1130
Db      |
Qy      297 CTGTGAGCGCACTAATATATCTATGAGAAACACAGACATCTGGATAAATAATCCATGACA 356
Db      |
Qy      1131 CTGTGAGCGCACTAATATATCTATGAGAAACACAGACATCTGGATAAATAATCCATGACA 1190
Db      |
Qy      357 TCTTTCATGTTTTTCCCAAGTCAAGAGGAAATGAGTTTATTTTGGCTCTGAAAT 416
Db      |
Qy      1191 TCTTTCATGTTTTTCCCAAGTCAAGAGGAAATGAGTTTATTTTGGCTCTGAAAT 1250
Db      |
Qy      417 GCAAAACAGGTTTCCGTCATTTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATATA 476
Db      |
Qy      1251 GCAAAACAGGTTTCCGTCATTTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATATA 1310
Db      |
Qy      477 AACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAGAGGAGA 536
Db      |
Qy      1311 AACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCTTATCAAGAGGAGA 1370
Db      |
Qy      537 TAGCAATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAAATATCCAAAGTTG 596
Db      |
Qy      1371 TAGCAATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAAATATCCAAAGTTG 1430
Db      |
Qy      597 ATGAAGTCAGAGGCTGGTATATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCAC 656
Db      |
Qy      1431 ATGAAGTCAGAGGCTGGTATATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCAC 1490
Db      |
Qy      657 TGTACGTAGTCAGTTACGTAATCTGGAGAGGTGACAGGCTGACCTGACCTGGTACT 716
Db      |
Qy      1491 TGTACGTAGTCAGTTACGTAATCTGGAGAGGTGACAGGCTGACCTGGTACT 1550
Db      |
Qy      717 CACATTCCTGTCGATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGA 776
Db      |
Qy      1551 CACATTCCTGTCGATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGA 1610
Db      |
Qy      777 AGAATCCACA CTGTGTGTCCTTTTACAGCTATCAAGTCTGTAAGATGACCCAACTTGA 836
Db      |
Qy      1611 AGAATCCACA CTGTGTGTCCTTTTACAGCTATCAAGTCTGTAAGATGACCCAACTTGA 1670
Db      |
Qy      837 AAACAAAGGAATTTTGGGCCCACTTTTGGATTCAGCAGGTCTCTTCTGCTGACTATCTC 896
Db      |
Qy      1671 AAACAAAGGAATTTTGGGCCCACTTTTGGATTCAGCAGGTCTCTTCTGCTGACTATCTC 1730
Db      |

897 CTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACA 956
1731 CTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACA 1790
957 AGCCTCATGATCTACAGCCTCGAAAGAAATATCTTACTGTGCTGTTCATATATATGCTG 1016
1791 AGCCTCATGATCTACAGCCTCGAAAGAAATATCTTACTGTGCTGTTCATATATATGCTG 1850
1017 CTC----- 1020
1851 CTCAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCGCTTGAATACCC 1910
1021 ----- 1020
1911 TAGCCTCTCTAGGTTATGTGTTAGTAGATAGACAA CAGGGGATCTGTCTCACCGAGGC 1970
1021 ----- 1020
1971 TTAAATTTGAAGGCGCTTTTAAATATATAAATAGGGTCAAAATAGAAATTGACGATCAGGTGG 2030
1021 ----- 1020
2031 AAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCC 2090
1021 ----- 1020
2091 ACGCTGGTCTTATGGAGGATACCTCTCCGTATGGCAATTAATGAGAGGTCAAGATATCT 2150
1021 ----- GGTTCGTATTGCTGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACA 1076
2151 TCAGGTTGTTGCTGTGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACA 2210
1077 CGAAAGCTTATATGGGTCAACCTGACCAAGAAATGAACAGGGCTATTTACTTGGATCTGTGG 1136
2211 CGAAAGCTTATATGGGTCAACCTGACCAAGAAATGAACAGGGCTATTTACTTGGATCTGTGG 2270
1137 CCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTTACTGCTCTTACATGTTTCC 1196
2271 CCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTTACTGCTCTTACATGTTTCC 2330
1197 TGGATGAGAAATGCTCAATTTTGACATACCAAGTATATTTACTGAGTTTTTGTAGTGGGCTG 1256
2331 TGGATGAGAAATGCTCAATTTTGACATACCAAGTATATTTACTGAGTTTTTGTAGTGGGCTG 2390
1257 GAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAAT 1316
2391 GAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTCTGAAT 2450
1317 CGGAGAAACATTTATGAACCTGTCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCACGTA 1376
2451 CGGAGAAACATTTATGAACCTGTCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCACGTA 2510
1377 TTGCTGCTCTAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTCGTATACACTGGCT 1436
2511 TTGCTGCTCTAAAGTGATATAAATTTTGACCTGTGTAGAACTCTCTCGTATACACTGGCT 2570
1437 ATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATTTGATCATCATNTTGA 1496
2571 ATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGAAATTTGATCATCATNTTGA 2630
1497 TACTCGCATGTAACATCTACTCTCTGAAAATAAATGTGGTGCCCATGACGGGCTACGGT 1556
2631 TACTCGCATGTAACATCTACTCTCTGAAAATAAATGTGGTGCCCATGACGGGCTACGGT 2690
1557 TTGTGTAGTAAATCTAATAACCTTAAACCCACATCTCAAAATCAAAATGATACATATTCCT 1616
2691 TTGTGTAGTAAATCTAATAACCTTAAACCCACATCTCAAAATCAAAATGATACATATTCCT 2750
1617 GAGAGACCCAGCAATACCATTAAGATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 1661
2751 GAGAGACCCAGCAATACCATTAAGATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 2795
```



RESULT 8	
US-09-976-674-2 ;	
; Sequence 2, Application US/09976674	
; Patent No. 6844180	
; GENERAL INFORMATION:	
; APPLICANT: Qi, Steve	
; APPLICANT: Akinsanya, Karen	
; APPLICANT: Riviere, Pierre	
; APPLICANT: Junien, Jean-Louis	
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV	
; FILE REFERENCE: 70669	
; CURRENT APPLICATION NUMBER: US/09/976,674	
; CURRENT FILING DATE: 2001-10-12	
; PRIOR APPLICATION NUMBER: US 60/240,117	
; PRIOR FILING DATE: 2000-10-12	
; NUMBER OF SEQ ID NOS: 61	
; SOFTWARE: PatentIn version 3.1	
; SEQ ID NO 2	
; LENGTH: 2671	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-09-976-674-2	
Query Match 65.1%; Score 1085.8; DB 3; Length 2671;	
Best Local Similarity 82.2%; Pred. No. 7.le-289;	
Matches 1397; Conservative 0; Mismatches 2; Indels 300; Gaps 1;	
Qy	1 AACAGGTACAGCAAAATCCTAAAGTGCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGA 60
Db	958 AACAGGTACAGCAAAATCCTAAAGTGCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGA 1017
Qy	61 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGAGATTCATTGTA 120
Db	1018 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGAGATTCATTGTA 1077
Qy	121 AGGAGTTGAATATATTGCGAGAGCTGGATGGATCTCTCAGGGAATAATATGCTTGGTCCAT 180
Db	1078 AGGAGTTGAATATATTGCGAGAGCTGGATGGATCTCTCAGGGAATAATATGCTTGGTCCAT 1137
Qy	181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCAATTAATTAT 240
Db	1138 CCTACTAGATCGCTCCAGACTCGCTCGAGATAGTGTGATCTCACTCAATTAATTAT 1197
Qy	241 CCAGTAGAAGATGATGTTATGGAAGGACAGACATCAATGAGTCAGTGCCTGATTCGT 300
Db	1198 CCAGTAGAAGATGATGTTATGGAAGGACAGACATCAATGAGTCAGTGCCTGATTCGT 1257
Qy	301 GAGCCCACTAATATCTATGAAAGAACACAGACATCTGGATAAATATCCATGACATCTT 360
Db	1258 GAGCCCACTAATATCTATGAAAGAACACAGACATCTGGATAAATATCCATGACATCTT 1317
Qy	361 TCATGTTTTTCCCAAGTCACGAGAGGAATTCAGTGTATTTTTTGGCTCTGAATGCAA 420
Db	1318 TCATGTTTTTCCCAAGTCACGAGAGGAATTCAGTGTATTTTTTGGCTCTGAATGCAA 1377
Qy	421 AACAGGTTTCCGTCATTATACAAAATTACATCTATTTTTTAAAGGAAAGCAAAATATAACG 480
Db	1378 AACAGGTTTCCGTCATTATACAAAATTACATCTATTTTTTAAAGGAAAGCAAAATATAACG 1437
Qy	481 ATCCAGTGGGGCTGCCTGCTCCAAAGTGAATTCAGTGTCTTCAAGAGGAGATAGC 540
Db	1438 ATCCAGTGGGGCTGCCTGCTCCAAAGTGAATTCAGTGTCTTCAAGAGGAGATAGC 1497
Qy	541 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 600
Db	1498 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 1557
Qy	601 AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTA 660
Db	1558 AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACCTGTA 1617
Qy	661 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGA CAAGGCTGACTGACCGTGGCTACTCACA 720



```
; Sequence 4, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-4

Query Match      52.8%; Score 881; DB 3; Length 1197;
Best Local Similarity 87.2%; Pred. No. 1,4e-232;
Matches 1044; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

QY 618 ATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAAGTTACGTAA 677
DB 1 ATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAAGTTACGTAA 60
QY 678 ATCTCGAGAGTGCAAGGCTGACTGACCGTGCTACTACATCTTGTGCTGATCAGTC 737
DB 61 ATCTCGAGAGTGCAAGGCTGACTGACCGTGCTACTACATCTTGTGCTGATCAGTC 120
QY 738 AGCACTGTGACTCTTTTAAGTAAGTATAGTAACAGAGAATCCACATGTGTGTC 797
DB 121 AGCACTGTGACTCTTTTAAGTAAGTATAGTAACAGAGAATCCACATGTGTGTC 180
QY 798 TTTAACAAGCTATCAAGTCTCGAAGATGACCAACTTGCACAAAGAGAAATTTTGGGCCA 857
DB 181 TTTAACAAGCTATCAAGTCTCGAAGATGACCAACTTGCACAAAGAGAAATTTTGGGCCA 240
QY 858 CCATTTTGGATTACAGAGTCTCTCTCTGACTATCTCTCCAGAGAAATTTCTCTTTTG 917
DB 241 CCATTTTGGATTACAGAGTCTCTCTCTGACTATCTCTCCAGAGAAATTTCTCTTTTG 300
QY 918 AAAGTACTACTGATTATCAATGTATGGATGCTCTACAAGCCTCATGATCTACAGCTG 977
DB 301 AAAGTACTACTGATTATCAATGTATGGATGCTCTACAAGCCTCATGATCTACAGCTG 360
QY 978 GAAAGAAATATCTACTGTGCTGTTTATATATATATGTTGGTCTCTCAGGGTCAAATAGAAATTG 1020
DB 361 GAAAGAAATATCTACTGTGCTGTTTATATATATATGTTGGTCTCTCAGGGTCAAATAGAAATTG 420
QY 1021 ----- 1020
DB 421 ACGATCAGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATC 480
QY 1021 ----- 1020
DB 481 GTGTGGGCATCCAGCGTGCTCTATGAGATACCTCTCCCTGATGGCATTAATCCAGA 540
QY 1021 ----- GGTGTCTATTGCTGGGGCCCAAGTCACTCTGTGATCTTCTATG 1064
DB 541 GGTGAGATATCTTCAGGGTGTCTATGCTGGGGCCCAAGTCACTCTGTGATCTTCTATG 600
QY 1065 ATACAGGATACACGGAACTTATATGTTGCTACCTGACCAAGATGAACAGGGCTATTACT 1124
DB 601 ATACAGGATACACGGAACTTATATGTTGCTACCTGACCAAGATGAACAGGGCTATTACT 660
QY 1125 TAGGATCTGTGCCATGCAAGCAGAAAAAGTTCCCTCTCTGAACCAAAATCGTTTACTGCTCT 1184
```

```
DB 661 TAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCT 720
QY 1185 TACATGTTTCTCGATGAGATGTCATTTTGCAATACCAAGTATATATCTGAGTTTTT 1244
DB 721 TACATGTTTCTCGATGAGATGTCATTTTGCAATACCAAGTATATATCTGAGTTTTT 780
QY 1245 TAGTGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAA 1304
DB 781 TAGTGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAA 840
QY 1305 GAGTTCTCTGAATCGGAGAACATTTATGAATCTCATCTTTTGCACTACCTTCAAGAAAAACC 1364
DB 841 GAGTTCTCTGAATCGGAGAACATTTATGAATCTCATCTTTTGCACTACCTTCAAGAAAAACC 900
QY 1365 TTGATCACGTATTGCTGCTTAAAGATGATATAATTTTGACCTGTAGAACTCTCTGG 1424
DB 901 TTGATCACGTATTGCTGCTTAAAGATGATATAATTTTGACCTGTAGAACTCTCTGG 960
QY 1425 TATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACACAGAAATTGATC 1484
DB 961 TATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACACAGAAATTGATC 1020
QY 1485 ATCACTTTTGTATACCTGCTGATGATCACTTCTCTGAAATTAATGTTGCTGATGCA 1544
DB 1021 ATCACTTTTGTATACCTGCTGATGATCACTTCTCTGAAATTAATGTTGCTGATGCA 1080
QY 1545 GGGGTCTACGGTTTGTGCTAGTAATCTATACCTTAAACCCACATGCTTCAAAATCAATG 1604
DB 1081 GGGGTCTACGGTTTGTGCTAGTAATCTATACCTTAAACCCACATGCTTCAAAATCAATG 1140
QY 1605 ATCACTATCTCTGAGAGACCCAGCAATACCATAGAAATTTCTTAAAAAATAAAAAA 1661
DB 1141 ATCACTATCTCTGAGAGACCCAGCAATACCATAGAAATTTCTTAAAAAATAAAAAA 1197

RESULT 10
US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 4309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-14

Query Match      45.8%; Score 765.2; DB 3; Length 4309;
Best Local Similarity 77.7%; Pred. No. 2e-200;
Matches 1082; Conservative 0; Mismatches 8; Indels 302; Gaps 1;

QY 579 GATCTAATATCCAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACT 638
DB 1222 GAAGGAAGATCCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACT 1281
QY 639 CCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGTGACAAAGC 698
DB 1282 CCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGTGACAAAGC 1341
QY 699 TGACTACCGTGGCTACTCACAATCTTCTGCTCAGTCAGTCAGTCTCTTTTATAA 758
DB 1342 TGACTACCGTGGCTACTCACAATCTTCTGCTCAGTCAGTCAGTCTCTTTTATAA 1401
```

Qy	759	GTAAAGTATAGTAACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCTG	818
Db	1402	GTAAAGTATAGTACACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCTG	1461
Qy	819	AAGATGACCCAACTTGCAAAAACAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTC	878
Db	1462	AAGATGACCCAACTTGCAAAAACAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTC	1521
Qy	879	CTCTTCCTGACTATATCTCTCAGAAATTTCTCTTTTGAAAGTACTACTGGATTTACAT	938
Db	1522	CTCTTCCTGACTATATCTCTCAGAAATTTCTCTTTTGAAAGTACTACTGGATTTACAT	1581
Qy	939	TGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGGAAGAAATATCTACTGTGC	998
Db	1582	TGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGGAAGAAATATCTACTGTGC	1641
Qy	999	TGTTTCATATATGGTGTCTC-----	1019
Db	1642	TGTTTCATATATGGTGTCTCCTCAGGTGAGTGTGTGATTAATCGGTTTAAAGAGTCAA	1701
Qy	1020	-----	1019
Db	1702	GTATTTCCGTTGAATACCTAGCCTCTCTAGGTTATGTGTTGTAGTGATAGACAACAG	1761
Qy	1020	-----	1019
Db	1762	GGGATCTGTACCGAGGCTTAAATTTGAAGGGCCCTTTTAAATATAAAATGGGTCAAAT	1821
Qy	1020	-----	1019
Db	1822	AGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGA	1881
Qy	1020	-----	1019
Db	1882	CTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGATACCTCTCCCTGATGGCAAT	1941
Qy	1020	-----AGTTTGTCTATGCTGGGCCCCAGTCACTCTGTGGAT	1056
Db	1942	AAATGCAAGGTCAATATCTCAGGGTTCCTATGCTGGGCCCCAGTCACTCTGTGGAT	2001
Qy	1057	CTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACAGCAATGAACAGGG	1116
Db	2002	CTTCTATGATACAGGATACACGGAACGTTATATGGGTCAACCTGACAGCAATGAACAGGG	2061
Qy	1117	CTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACAAATCGTTT	1176
Db	2062	CTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACAAATCGTTT	2121
Qy	1177	ACTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGCACATACCAATATTTACT	1236
Db	2122	ACTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGCACATACCAATATTTACT	2181
Qy	1237	GAGTTTATAGTACAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGAGACA	1296
Db	2182	GAGTTTATAGTACAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGAGACA	2241
Qy	1297	CAGCATAGAGTTCTCGAATCGGAGAACATTTAGAACTGCATCTTTTGCACCTACCTTCA	1356
Db	2242	CAGCATAGAGTTCTCGAATCGGAGAACATTTAGAACTGCATCTTTTGCACCTACCTTCA	2301
Qy	1357	AGAAAACCTTGGATCAGTATGCTGCTCTAAAGTGATATAAATTTTGAACCTGTAGAA	1416
Db	2302	AGAAAACCTTGGATCAGTATGCTGCTCTAAAGTGATATAAATTTTGAACCTGTAGAA	2361
Qy	1417	CTCTCTGGTATACATGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAAACACAG	1476
Db	2362	CTCTCTGGTATACATGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAAACACAG	2421
Qy	1477	AATTGATCATCATTTTGAATCTGCGATGTAAACATCTACTCTCTGAAATTAATGTGGT	1536
Db	2422	AATTGATCATCATTTTGAATCTGCGATGTAAACATCTACTCTCTGAAATTAATGTGGT	2481

Qy	1537	GCCATGCGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCAATGCTCAAA	1596
Db	2482	GCCATGCGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCCAATGCTCAAA	2541
Qy	1597	ATCAATGATACATATTTCTTGAGACCCAGCAATACCATTAAGATTTACTTAAAAAANA	1656
Db	2542	ATCAATGATACATATTTCTTGAGACCCAGCAATACCATTAAGATTTACTTAAAAAANA	2601
Qy	1657	AAAAAAAAAAAAA	1668
Db	2602	AAAAAAAAAAAAA	2613
RESULT 11			
US-09-280-116-171			
; Sequence 171, Application US/09280116A			
; Patent No. 6331427			
; GENERAL INFORMATION:			
; APPLICANT: Robison, Keith E.			
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs			
; FILE REFERENCE: 5800-24, 035800/176965			
; CURRENT APPLICATION NUMBER: US/09/280,116A			
; CURRENT FILING DATE: 1999-03-26			
; NUMBER OF SEQ ID NOS: 268			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 171			
; LENGTH: 823			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION: prollyl oligopeptidases			
US-09-280-116-171			
Query Match 29.6%; Score 493.4; DB 3; Length 823;			
Best Local Similarity 93.1%; Pred. No. 8.1e-126;			
Matches 625; Conservative 0; Mismatches 11; Indels 35; Gaps 9;			
Qy	1021	GGTTGCTATTGCTGGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	1080
Db	153	GGTTGCTATTGCTGGGTCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	212
Qy	1081	ACGTTATATGGGTCAACCTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCAT	1140
Db	213	ACGTTATATGGGTCAACCTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCAT	272
Qy	1141	GCAAGCAGAAAAGTTCCCTCTGAAACCAATGTTTACTGTCTTCTTACATGCTTCTCTGGA	1200
Db	273	GCAAGCAGAAAAGTTCCCTCTGAAACCAATGTTTACTGTCTTCTTACATGCTTCTCTGGA	332
Qy	1201	TGGAATGTCCATTTTGGCAGATATATTTACTGAGTTTGTAGTGAGGCTGGAAA	1260
Db	333	TGGAATGTCCATTTTGGCAGATATATTTACTGAGTTTGTAGTGAGGCTGGAAA	392
Qy	1261	GCCATATGATTTTAC-----AGATCTATCTCTCA--GGAGAG	1293
Db	393	GCAGTATGATTTACAGTATCTTTATTTTGTGTGTGTAAGATCTATCTCAGGAGAG	452
Qy	1294	ACA--CAGCATAGAGTTCC--TGAATCGGAGAGAACATTATGAATCTTTT--GCAT	1349
Db	453	ACACCGCATAGAGTTCCCTGAATCGGGGAGAGAACATTATGAATCTTTTGGCACT	512
Qy	1350	ACCTTCAAG--AAAACTTGGATCA--CGTATGCTGCT--CTAAAAATGATTAATTTGAC	1406
Db	513	ACCTTCAAGAAAAACCTTGGATCACCTTATGCTGCTCTTAAATGATTAATTTGAC	572
Qy	1407	CTGTGTAGAACTCTCT--GGTATACCTGGCTATTTTAAACCAATGAGGAGTTTAAATCAAC	1465
Db	573	CTGTGTAGAACTCTCTGGGTTATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATCAAC	632
Qy	1466	AGAAAACACAGAAATGATCATCATCTTTTGGATACCTGCGCATGTAACATCTACTCTCGAAA	1525
Db	633	AGAAAACACAGAAATGATCATCATCTTTTGGATACCTGCGCATGTAACATCTACTCTCGAAA	692





; SEQ ID NO 30		
; LENGTH: 4159		
; TYPE: DNA		
; ORGANISM: Homo sapiens		
us-09-976-674-30		
Query Match		
Best Local Similarity 57.3%; Pred. No. 8.2e-76;		
Matches 587; Conservative 0; Mismatches 435; Indels 3; Gaps 1;		
Qy	2	ACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTGAGAAATAATGATGTGCTGAA 61
Db	1288	ACAGGCGACGAAGATCCCAAGATTGCTTTGAACCTGGCTGAGTTCAGACTGCAGCCAG 1347
Qy	62	GGAAGGATCATAGATGTCTAGATAAGGAACCTAATTTCAACCTTTTGAGATTCTATTGAA 121
Db	1348	GGCAAGATCTCTCGACCCCGAGAGAGAGCTGGTGCAGCCCTTCAGCTGGCTGTTCCCG 1407
Qy	122	GGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGAGAAATATGCTTGGTCCATC 181
Db	1408	AAGGTGGAGTACATCGCCAGGGCCGGGTGGACCCGGGATGGCAAAATACGCTGGGCCATG 1467
Qy	182	CTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATTATTTATC 241
Db	1468	TTCTGAGCCGGCCCCAGCAGTGGCTCCAGCTCGTCTCTCCGCCCGCCCTGTTTCATC 1527
Qy	242	CCAGTAGAAGATGATGTATTGAAAGGCGAGAGACTCATTGAGTCAGTGCCTGATTTCTGTG 301
Db	1528	CCGAGCACAGAAATGAGGAGCAGCGGCTAGCCTCTGCCAGAGCTGTCCCAGGATGTC 1587
Qy	302	ACGCCACTAATTATCTATGAGAAACACAGACATCTGGATATAATATCCATGACATCTTT 361
Db	1588	CAGCGGTATGTGTGTACGAGGAGGTCCCAACGTCTGANTCAATGTTCAIGACATCTTC 1647
Qy	362	CATGTTTTTCCCAAGTCAAGAG---GAATTGAGTTTATTTTGGCTCTGAATGC 418
Db	1648	TATCCCTTCCCCCAATCAGGAGAGGACGAGCTCTGCTTCTCCCGCCCAATGAATGC 1707
Qy	419	AAACAGGTTTCGTCATTATACAAAATTACATCTATTTAAAGGAAAGCAAAATATAA 478
Db	1708	AAGACGGCTTCTGCGCATTTGTACAAAGTCAACCGCGTTTAAAAATCCAGGGGCTACGAT 1767
Qy	479	CGATCCAGTGGTGGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAAGAGGAGATA 538
Db	1768	TGGAGTGAGCCCTTCAGCCCCGGGGAAGATGAATTTAAGTGCCCATTAAGGAAGAGATT 1827
Qy	539	GCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAGTTGAT 598
Db	1828	GCTCTGACCAAGCGGTGAATGGGAGGTTTTTGGCGAGGCACGGCTCCAAGATCTGGGTCAAT 1887
Qy	599	GAAGTCAGAAAGGCTGGTATATTTGAAGGCACCAAGACTCCCTCTTTAGAGCATCACTG 658
Db	1888	GAGGAGACCAAGCTGGTGACTTCCAGGGCACCAAGGACACGCGCTGGAGCACCACTC 1947
Qy	659	TACGTAGTCAGTTTACGTAAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCA 718
Db	1948	TACGTGTCAGCTATGAGGGGCGCGCGAGATCGTACGCTTCAACACGCCGCGCTTCTCC 2007
Qy	719	CATTTCTTGTGTCATCAGTCAGCATGTGACTTCTTTTATAAGTAAGTATAGTAACCAAG 778
Db	2008	CATAGCTGCTCCATGAGCCAGAACTTCGACATGTTTCGTACGCCACTACAGCAGGTGAGC 2067
Qy	779	AATCCACATGTGTGTCCTTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAA 838
Db	2068	ACGCCGCCCTGGGTGACGCTTACAGCTGAGGGGCCCGAGCAGACCCCTCGACACAG 2127
Qy	839	ACAAGGAATTTTGGGGCCACATTTTGGATTGAGAGTCTCTTCTCTGACTATCTCTCT 898
Db	2128	CAGCCCCGCTTCTGGGCTAGCATGTGAGGCGAGCGAGCTGCCCCCGGATTTATGTTCT 2187
Qy	899	CCAGAAATTTTCTTTTGAAGTACTACTGGAATTAACATTTGATGGGATGCTCTACAAG 958
Db	2188	CCAGAGATCTTCCATTTCCACACGCGCTCGGATGTGCGGCTCTACGGCATGATCTACAAG 2247

Qy	959	CCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTCTGTTCATATATGGTGGTCTT 1018
Db	2248	CCCCACGCCTTGCAGCCAGGGAAGACACCCACCCCTTTGTATATGGAGGGCCC 2307
Qy	1019	CAGGT 1023
Db	2308	CAGGT 2312

Search completed: May 3, 2006, 02:12:47  
Job time : 547.632 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

QM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:41:02 ; Search time 1484.73 Seconds  
(without alignments)  
9295.694 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 1669  
Sequence: 1 acacagtcacgaatccta.....aaaaaaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134699005 residues  
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubna/US09A\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1669	100.0	1669	8	US-10-825-632-6
2	1667	99.9	2830	9	US-10-956-157-2177
3	1648.4	98.8	4523	3	US-09-976-674-8
4	1648.4	98.8	4523	9	US-10-982-512-8
5	1501	89.9	4676	3	US-09-976-674-20
6	1501	89.9	4676	9	US-10-982-512-20
7	1353.4	81.1	3143	6	US-10-170-789-37
8	1349.6	80.9	4829	3	US-09-976-674-12
9	1349.6	80.9	4829	9	US-10-982-512-12
10	1347	80.7	3120	7	US-10-415-122-5
11	1347	80.7	3120	8	US-10-825-632-2
12	1344	80.5	3106	7	US-10-311-035-30
13	1343.6	80.5	4685	3	US-09-976-674-22
14	1343.6	80.5	4685	9	US-10-982-512-22
15	1250.6	74.9	2510	7	US-10-275-505-16
16	1250.6	74.9	2510	10	US-11-140-224-16
17	1087.4	65.2	2649	6	US-10-170-789-39
18	1085.8	65.1	2649	3	US-10-054-776-1
19	1085.8	65.1	2671	3	US-09-976-674-2
20	1085.8	65.1	2671	9	US-10-982-512-2
21	881	52.8	1197	8	US-10-825-632-4
22	765.2	45.8	4309	3	US-09-976-674-14
23	765.2	45.8	4309	9	US-10-982-512-14

24	636	38.1	925	6	US-10-264-237-710	Sequence 710, Appl
25	600	35.9	600	9	US-10-956-157-7412	Sequence 7412, Ap
26	434.2	26.0	1083	8	US-10-825-632-8	Sequence 8, Appli
27	321	19.2	3287	7	US-10-415-122-3	Sequence 3, Appli
28	313	18.8	2495	7	US-10-415-122-8	Sequence 8, Appli
29	313	18.8	2617	3	US-09-976-674-4	Sequence 4, Appli
30	313	18.8	2617	9	US-10-982-512-4	Sequence 4, Appli
31	313	18.8	2660	7	US-10-072-012-223	Sequence 223, App
32	313	18.8	2660	9	US-10-072-012-225	Sequence 225, App
33	313	18.8	3000	7	US-10-415-122-1	Sequence 1, Appli
34	313	18.8	3716	9	US-10-433-757-30	Sequence 30, Appl
35	313	18.8	4076	3	US-09-976-674-32	Sequence 32, Appl
36	313	18.8	4076	9	US-10-982-512-32	Sequence 32, Appl
37	313	18.8	4159	3	US-09-976-674-30	Sequence 30, Appl
38	313	18.8	4159	9	US-10-982-512-30	Sequence 30, Appl
39	313	18.8	4219	3	US-09-976-674-28	Sequence 28, Appl
40	313	18.8	4219	9	US-10-982-512-28	Sequence 28, Appl
41	313	18.8	4302	3	US-09-976-674-24	Sequence 24, Appl
42	313	18.8	4302	9	US-10-982-512-24	Sequence 24, Appl
43	255	15.3	2261	6	US-10-094-749-47	Sequence 47, Appl
44	250.6	15.0	4037	3	US-09-976-674-40	Sequence 40, Appl
45	250.6	15.0	4037	9	US-10-982-512-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1  
US-10-825-632-6  
; Sequence 6, Application US/10825632  
; Publication No. US20040191826A1  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: CORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: FCSB-100-Div. 1  
; CURRENT APPLICATION NUMBER: US/10/825,632  
; CURRENT FILING DATE: 2004-04-15  
; PRIOR APPLICATION NUMBER: US 10/070,464  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1669  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-825-632-6

Query Match	100.0%;	Score 1669;	DB 8;	Length 1669;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 1669;	Conservative 0;			
Qy	1	AACAGGTACAGCAAACTCTAAAGTACACTTTTAAAGATGTCAGAAATAATGATTGATCTGA	60	
Db	1	AACAGGTACAGCAAACTCTAAAGTACACTTTTAAAGATGTCAGAAATAATGATTGATCTGA	60	
Qy	61	AGGAAGATCATAGATGTCATAGATAAGCACTAATTCACCTTTTGAGATTCATTGGA	120	
Db	61	AGGAAGATCATAGATGTCATAGATAAGCACTAATTCACCTTTTGAGATTCATTGGA	120	
Qy	121	AGGAGTTGAATATATATGCGCAGAGCTGGATGACTCCTCGAGGAAATAATGCTGGTCCAT	180	
Db	121	AGGAGTTGAATATATATGCGCAGAGCTGGATGACTCCTCGAGGAAATAATGCTGGTCCAT	180	
Qy	181	CCTACTAGATCGTCCAGACTCGCCCTACAGATAGTGTGATCTCACCTGAATTTAT	240	
Db	181	CCTACTAGATCGTCCAGACTCGCCCTACAGATAGTGTGATCTCACCTGAATTTAT	240	

Qy 241 CCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCATTGAGTCAGTGCCTGATCTGT 300  
Db 241 CCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCATTGAGTCAGTGCCTGATCTGT 300  
Qy 301 GAGCCCACTAATTTATCTATGAGAACACACAGACATCTGGATTAATATCCATGACATCTT 360  
Db 301 GAGCCCACTAATTTATCTATGAGAACACACAGACATCTGGATTAATATCCATGACATCTT 360  
Qy 361 TCATGTTTTCCCAAGTCAGAACAGGAAATGAGTTTATTTTAAAGGAAAGCAAAATATAACG 420  
Db 361 TCATGTTTTCCCAAGTCAGAACAGGAAATGAGTTTATTTTAAAGGAAAGCAAAATATAACG 420  
Qy 421 AACAGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAACG 480  
Db 421 AACAGTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAACG 480  
Qy 481 ATCCAGTGTGGGCTGCCCTGCCAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGC 540  
Db 481 ATCCAGTGTGGGCTGCCCTGCCAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGC 540  
Qy 541 AATTACAGTGTGTAATGGAAAGTTCCTTGGCGGCATGGATCTAATATCCAAGTTGATGA 600  
Db 541 AATTACAGTGTGTAATGGAAAGTTCCTTGGCGGCATGGATCTAATATCCAAGTTGATGA 600  
Qy 601 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTA 660  
Db 601 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTA 660  
Qy 661 CGTAGTCAGTACGTAAATCTCGAGAGGTCGACAGGCTGACTGACCGTGGCTACTCACA 720  
Db 661 CGTAGTCAGTACGTAAATCTCGAGAGGTCGACAGGCTGACTGACCGTGGCTACTCACA 720  
Qy 721 TTCTTCTGCATCAGTCAGCTGACTGCTCTTTAAGTAAGTATAGTATGACCAAGAA 780  
Db 721 TTCTTCTGCATCAGTCAGCTGACTGCTCTTTAAGTAAGTATAGTATGACCAAGAA 780  
Qy 781 TCACACTGTGTGCTCTTACAGCTATCAAGTCTCTGAAGTGAACCCAACTTGCAAAAC 840  
Db 781 TCACACTGTGTGCTCTTACAGCTATCAAGTCTCTGAAGTGAACCCAACTTGCAAAAC 840  
Qy 841 AAGGAATTTTGGGCCACATTTGGATTACAGAGGTCCTCTTCTGACTATATCCTCC 900  
Db 841 AAGGAATTTTGGGCCACATTTGGATTACAGAGGTCCTCTTCTGACTATATCCTCC 900  
Qy 901 AGAAATTTCTCTTTGAAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCC 960  
Db 901 AGAAATTTCTCTTTGAAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCC 960  
Qy 961 TCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTATATATGTTGTTCTCA 1020  
Db 961 TCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTATATATGTTGTTCTCA 1020  
Qy 1021 GGTGTGCTATTGTGGGGCCCAAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGA 1080  
Db 1021 GGTGTGCTATTGTGGGGCCCAAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGA 1080  
Qy 1081 AGTTATATGGGTACCCCTGACCAAGATGACAGGGCTATTACTAGGATCTGTGGCAT 1140  
Db 1081 AGTTATATGGGTACCCCTGACCAAGATGACAGGGCTATTACTAGGATCTGTGGCAT 1140  
Qy 1141 GCAGCAGAAAGTCCCTCTGAACCAATCGTTTACTGCTCTTACATGAGTTTCTGGA 1200  
Db 1141 GCAGCAGAAAGTCCCTCTGAACCAATCGTTTACTGCTCTTACATGAGTTTCTGGA 1200  
Qy 1201 TGAGATGTCATTTTGCACATACAGTATATTAAGTATTTTATGAGGGCTGGAAA 1260  
Db 1201 TGAGATGTCATTTTGCACATACAGTATATTAAGTATTTTATGAGGGCTGGAAA 1260  
Qy 1261 GCCATATGATTTACAGATCTATCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGG 1320  
Db 1261 GCCATATGATTTACAGATCTATCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGG 1320

Qy 1321 AGACATTATGAACCTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTATTGC 1380  
Db 1321 AGACATTATGAACCTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTATTGC 1380  
Qy 1381 TGCTCTAAAAGTGATATAATTTTACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440  
Db 1381 TGCTCTAAAAGTGATATAATTTTACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440  
Qy 1441 AACCAATAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATCTTTTGATACC 1500  
Db 1441 AACCAATAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATCTTTTGATACC 1500  
Qy 1501 TGCCATGTAACTACTCTCTGAAATATAATGTGGTGCATGCGGGGTCTACGGTTGT 1560  
Db 1501 TGCCATGTAACTACTCTCTGAAATATAATGTGGTGCATGCGGGGTCTACGGTTGT 1560  
Qy 1561 GGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGA 1620  
Db 1561 GGTAGTAATCTAATACCTTTAACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGA 1620  
Qy 1621 GACCCAGCAATACCAATGAATTTACTAAAAAAGGAAAAAAGAAAAA 1669  
Db 1621 GACCCAGCAATACCAATGAATTTACTAAAAAAGGAAAAAAGAAAAA 1669

RESULT 2

US-10-956-157-2177  
; Sequence 2177, Application US/10956157  
; Publication No. US20050118625A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth, William  
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH  
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES  
; FILE REFERENCE: 031996-043000 (AM 101081)  
; CURRENT APPLICATION NUMBER: US/10/956,157  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 319805  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2177  
; LENGTH: 2830  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-956-157-2177

Query Match 99.9%; Score 1667; DB 9; Length 2830;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAAATATGATGATGCTGA 60  
Db 1164 AACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAAATATGATGATGCTGA 1223  
Qy 61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTTCAGATTTCTATTGA 120  
Db 1224 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTTCAGATTTCTATTGA 1283  
Qy 121 AGGAGTTGAATATATGTCAGAGCTGGATGATCTCTGAGGAAAAATATCTGGTCCAT 180  
Db 1284 AGGAGTTGAATATATGTCAGAGCTGGATGATCTCTGAGGAAAAATATCTGGTCCAT 1343  
Qy 181 CCTACTAGATCCGCTCCCAAGATCGGCTACAGATAGTGTGATCTCACCTGAAATATTTAT 240  
Db 1344 CCTACTAGATCCGCTCCCAAGATCGGCTACAGATAGTGTGATCTCACCTGAAATATTTAT 1403  
Qy 241 CCAGTGAAGATGATGTTATGGAAGGACGAGACTCATTTGAGTCAGTGCCTGATCTGT 300  
Db 1404 CCAGTGAAGATGATGTTATGGAAGGACGAGACTCATTTGAGTCAGTGCCTGATCTGT 1463  
Qy 301 GAGCCCACTAATTTATCTATGAGAAACACACAGACATCTGGATAAATATCCATGACATCTT 360  
Db 1464 GAGCCCACTAATTTATCTATGAGAAACACACAGACATCTGGATAAATATCCATGACATCTT 1523



QY 361 TCATGTTTTTCCCAAGTCAGAGAGGAAATTTGAGTTTATTTTTTGGCTCTCGAATGCAA 420  
Db 1524 TCATGTTTTTCCCAAGTCAGAGAGGAAATTTGAGTTTATTTTTTGGCTCTCGAATGCAA 1583  
QY 421 AACAGGTTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACG 480  
Db 1584 AACAGGTTTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACG 1643  
QY 481 ATCCAGTGTGGGCTCGCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAGAGAGAGATAGC 540  
Db 1644 ATCCAGTGTGGGCTCGCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAGAGAGAGATAGC 1703  
QY 541 AATACCAAGTGTGAATGGAGAGTTCTTGGCGGCATGGATCTAATATCAAGTTGATGA 600  
Db 1704 AATACCAAGTGTGAATGGAGAGTTCTTGGCGGCATGGATCTAATATCAAGTTGATGA 1763  
QY 601 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTTAGAGCATCACCTGTA 660  
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTTAGAGCATCACCTGTA 1823  
QY 661 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGTGACCTGACCGTGCTACTCACA 720  
Db 1824 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGTGACCTGACCGTGCTACTCACA 1883  
QY 721 TTCTTGCTGCATCAGTCAGCAGCTGACCTCTTTTATAGTAAGTATAGTAACCAAGAA 780  
Db 1884 TTCTTGCTGCATCAGTCAGCAGCTGACCTCTTTTATAGTAAGTATAGTAACCAAGAA 1943  
QY 781 TCCACACTGTGTGCTCCCTTTCAAGCTATCAAGTCTTGAAGTAGCACCCTTTGCAAAAC 840  
Db 1944 TCCACACTGTGTGCTCCCTTTCAAGCTATCAAGTCTTGAAGTAGCACCCTTTGCAAAAC 2003  
QY 841 AAAGGAATTTGGGCGCACATTTGGATTTCAGAGGTCTCTCTGACTATATCTCTCC 900  
Db 2004 AAAGGAATTTGGGCGCACATTTGGATTTCAGAGGTCTCTCTGACTATATCTCTCC 2063  
QY 901 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGAGTCTCTACAGGC 960  
Db 2064 AGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGAGTCTCTACAGGC 2123  
QY 961 TCATGATCTACAGCTGGAAAGAAATATCTCTGCTGTCTGTCTCATATATGCTGCTCA 1020  
Db 2124 TCATGATCTACAGCTGGAAAGAAATATCTCTGCTGTCTGTCTCATATATGCTGCTCA 2183  
QY 1021 GGTGCTATGCTGGGCGCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGA 1080  
Db 2184 GGTGCTATGCTGGGCGCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGA 2243  
QY 1081 ACCTTATATGGGTCAACCTGACCAAGATGAACAGGCTATTTACTTAGGATCTGTGGCCAT 1140  
Db 2244 ACCTTATATGGGTCAACCTGACCAAGATGAACAGGCTATTTACTTAGGATCTGTGGCCAT 2303  
QY 1141 GCAGCAGAAAAGTTCCCTCTGAACCAATTCGTTTACTGCTCTTACATGTTTCTCTGGA 1200  
Db 2304 GCAGCAGAAAAGTTCCCTCTGAACCAATTCGTTTACTGCTCTTACATGTTTCTCTGGA 2363  
QY 1201 TGAGATGTCCTATTTTGACATACCAAGTATATTTACTGATTTTTTTAGTAGGGCTGAAA 1260  
Db 2364 TGAGATGTCCTATTTTGACATACCAAGTATATTTACTGATTTTTTTAGTAGGGCTGAAA 2423  
QY 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACAGCATAGAGTTCTCTGAATCGGG 1320  
Db 2424 GCCATATGATTTACAGATCTATCTCTCAGGAGAGACAGCATAGAGTTCTCTGAATCGGG 2483  
QY 1321 AGAACATTTAGATGCTGATCTTTTGCACTACCTTCAAGAAAACCTTTGGATCAGTATTCG 1380  
Db 2484 AGAACATTTAGATGCTGATCTTTTGCACTACCTTCAAGAAAACCTTTGGATCAGTATTCG 2543  
QY 1381 TGCCTCTAAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGCTATACACTGGCTATTT 1440  
Db 2544 TGCCTCTAAAAGTGATATAATTTTGAACCTGTGTAGAACTCTCTGCTATACACTGGCTATTT 2603  
QY 1441 AACCAATGAGGAGGTTTTTAATCAACAGAAAAACAGAAATTTGATCATCATCTTTTGATACC 1500

Db 2604 AACCAATGAGGAGGTTTTTAATCAACAGAAAAACAGAAATTTGATCATCATTTTGATACC 2663  
QY 1501 TGGCATGTAACTACTACTCTCTGAAATAAATGCTGCTGCATGCGAGGGGTCTACCGTTTGT 1560  
Db 2664 TGGCATGTAACTACTACTCTCTGAAATAAATGCTGCTGCATGCGAGGGGTCTACCGTTTGT 2723  
QY 1561 GGTAGTAATCTAATATCTTAACTTAAACCCACATGCTCAAAATCAAAATGATATTCCTGAGA 1620  
Db 2724 GGTAGTAATCTAATATCTTAACTTAAACCCACATGCTCAAAATCAAAATGATATTCCTGAGA 2783  
QY 1621 GACCCAGCAATCCATAAGAAATTTACTTAAAAAAAATTTTTTTTTT 1667  
Db 2784 GACCCAGCAATCCATAAGAAATTTACTTAAAAAAAATTTTTTTTTT 2830  
RESULT 3  
US-09-976-674-8  
; Sequence 8, Application US/09976674  
; Patent No. US20020115943A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 4523  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-8  
Query Match 98.8%; Score 1648.4; DB 3; Length 4523;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;  
QY 1 AACAGGTACAGCAAAATCCCTAAAGTCACATTTTAAAGATGTCAGAAATATATGATGCTGA 60  
Db 1164 AACAGGTACAGCAAAATCCCTAAAGTCACATTTTAAAGATGTCAGAAATATATGATGCTGA 1223  
QY 61 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCTATTGA 120  
Db 1224 AGGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCTATTGA 1283  
QY 121 AGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGAATATGCTGGTCCAT 180  
Db 1284 AGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTCTGAGGGAATATGCTGGTCCAT 1343  
QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTAT 240  
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTAT 1403  
QY 241 CCAGTAGAAGATGATGTTATGGAAAGGAGAGACTCATTTGAGTCAGTCGCTGATCTGT 300  
Db 1404 CCAGTAGAAGATGATGTTATGGAAAGGAGAGACTCATTTGAGTCAGTCGCTGATCTGT 1463  
QY 301 GAGCCACATAATTTCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 360  
Db 1464 GAGCCACATAATTTCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTT 1523  
QY 361 TCATGTTTTTCCCAAGTCAGAGAGGAAATTTGAGTTTATTTTTTGGCTCTCGAATGCAA 420  
Db 1524 TCATGTTTTTCCCAAGTCAGAGAGGAAATTTGAGTTTATTTTTTGGCTCTCGAATGCAA 1583  
QY 421 AACAGGTTTCGTCATTTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACG 480

Db 1584 AACAGGTTCCGTCATTATACAAATTAACATCTATTATTAAGGAAACCAATATAAAG 1643  
Qy 481 ATCAAGTGGGCTGCCGTCCCAAGTGATTTCAAGTGTCCTATCAAGAGGAGATAGC 540  
Db 1644 ATCCAGTGGTGGGCTGCCGTCCCAAGTGATTTCAAGTGTCCTATCAAGAGGAGATAGC 1703  
Qy 541 AATTACAGTGGTGAATGGGAAGTTCTTGGCGGGCATGGATCTAATATCCAAGTTGATGA 600  
Db 1704 AATTACAGTGGTGAATGGGAAGTTCTTGGCGGGCATGGATCTAATATCCAAGTTGATGA 1763  
Qy 601 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTA 660  
Db 1764 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTTAGAGCATCACCTGTA 1823  
Qy 661 CGTAGTCAGTTACGTAATCTCGGAGAGGTGACAAGGCTGACGTCGCTGCTACTCACA 720  
Db 1824 CGTAGTCAGTTACGTAATCTCGGAGAGGTGACAAGGCTGACGTCGCTGCTACTCACA 1883  
Qy 721 TTCTTGCTGATCAGTCAGCACTGACTCTCTTATTAAGTAAGTATAGTATACCAAGAA 780  
Db 1884 TTCTTGCTGATCAGTCAGCACTGACTCTCTTATTAAGTAAGTATAGTATACCAAGAA 1943  
Qy 781 TCCACACTGTGTGCCCTTACAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAAC 840  
Db 1944 TCCACACTGTGTGCCCTTACAGCTATCAAGTCTCGAAGATGACCCAACTTGCAAAAC 2003  
Qy 841 RAAGGAATTTGGGCGACATTTTGGATTCAGCAGGTCCTCTTCCTGACTACTCCTCC 900  
Db 2004 RAAGGAATTTGGGCGACATTTTGGATTCAGCAGGTCCTCTTCCTGACTACTCCTCC 2063  
Qy 901 AGAAATTTCTTTTGAAGTACTACTGGAATTACATGTAATGGGATGCTCTACAAGCC 960  
Db 2064 AGAAATTTCTTTTGAAGTACTACTGGAATTACATGTAATGGGATGCTCTACAAGCC 2123  
Qy 961 TCATGATCTACAGCTCGAAAGAAATATCCTACTGTGCTGTTTCATATATGGTGGTCTCA 1020  
Db 2124 TCATGATCTACAGCTCGAAAGAAATATCCTACTGTGCTGTTTCATATATGGTGGTCTCA 2179  
Qy 1021 GGTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1080  
Db 2180 GGTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 2239  
Qy 1081 ACGTTATATGGGTCAACCTGACACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCAT 2299  
Db 2240 ACGTTATATGGGTCAACCTGACACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCAT 2359  
Qy 1141 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTCCTGGA 1200  
Db 2300 GCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTCCTGGA 2359  
Qy 1201 TGAGAAATGTCATTTTGCATACAGTATATTAAGTATTTTGTAGTGGGCTGGAAA 1260  
Db 2360 TGAGAAATGTCATTTTGCATACAGTATATTAAGTATTTTGTAGTGGGCTGGAAA 2419  
Qy 1261 GCCATATGATTTACAGATCTATCCTCAGGAGACACAGCATPAAGATTCCTGAATCGGG 1320  
Db 2420 GCCATATGATTTACAGATCTATCCTCAGGAGACACAGCATPAAGATTCCTGAATCGGG 2479  
Qy 1321 AGACATATGAATGCTGCTTTTGGACCTTACCTTGAAGAAACCTTGGATCACGATATGC 1380  
Db 2480 AGACATATGAATGCTGCTTTTGGACCTTACCTTGAAGAAACCTTGGATCACGATATGC 2539  
Qy 1381 TGCTCTAAAGTGAATTAATTTTGAACCTGTGTAGAACTCTCTGGTATACATGCTATTT 1440  
Db 2540 TGCTCTAAAGTGAATTAATTTTGAACCTGTGTAGAACTCTCTGGTATACATGCTATTT 2599  
Qy 1441 ACCCAATAGAGGTTTAAATCAACAGAAAAACAGAAATGATCATCACATTTTGTATACC 1500  
Db 2600 ACCCAATAGAGGTTTAAATCAACAGAAAAACAGAAATGATCATCACATTTTGTATACC 2659  
Qy 1501 TGGCATGTAAATCTACTCTCTGAAATTAATGTGGTGCATGCAGGGGTCTACGGTTTGT 1560  
Db 2660 TGGCATGTAAATCTACTCTCTGAAATTAATGTGGTGCATGCAGGGGTCTACGGTTTGT 2719

Qy 1561 GGTAGTAATCTAATACCTTAACCCCACTGCTCAAAATCAAAATGATACATATTCCTGAGA 1620  
Db 2720 GGTAGTAATCTAATACCTTAACCCCACTGCTCAAAATCAAAATGATACATATTCCTGAGA 2779  
Qy 1621 GACCCAGCAATACCATAGAAATTAATAAAAAAAAAAAAAAAAAAAAAA 1668  
Db 2780 GACCCAGCAATACCATAGAAATTAATAAAAAAAAAAAAAAAAAAAAAA 2827

RESULT 4  
US-10-982-512-8  
; Sequence 8, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; PRIOR FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 4523  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-982-512-8

Query Match 98.8%; Score 1648.4; DB 9; Length 4523;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1663; Conservative 0; Mismatches 1; Indels 4; Gaps 1;  
Qy 1 AACAGGTACACAAATCTTAAAGTCACCTTTTAAGATGTCGAAATTAATGATGATGCTGA 60  
Db 1164 AACAGGTACACAAATCTTAAAGTCACCTTTTAAGATGTCGAAATTAATGATGATGCTGA 1223  
Qy 61 AGGAAGGATCATAGATGTCTATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTTGA 120  
Db 1224 AGGAAGGATCATAGATGTCTATAGATAAGGAACCTAATTCACCTTTTGAGATTTCTATTTGA 1283  
Qy 121 AGGATTTGAATATATTCGACAGCTGGATGGACTCCTCGAGGGAAAAATATGCTTGGTCCAT 180  
Db 1284 AGGATTTGAATATATTCGACAGCTGGATGGACTCCTCGAGGGAAAAATATGCTTGGTCCAT 1343  
Qy 181 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTAT 240  
Db 1344 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTAT 1403  
Qy 241 CCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTGCCTGATTCCT 300  
Db 1404 CCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTGCCTGATTCCT 1463  
Qy 301 GACGCCACTAATTAATCTATGAAGAAACACAGACATCTGGATAAAATATCCATGACATCTT 360  
Db 1464 GACGCCACTAATTAATCTATGAAGAAACACAGACATCTGGATAAAATATCCATGACATCTT 1523  
Qy 361 TCATGTTTTTCCCCAAAGTCACGAGAGAAATTTGAGTTTATTTTGGCTCTGATGCAA 420  
Db 1524 TCATGTTTTTCCCCAAAGTCACGAGAGAAATTTGAGTTTATTTTGGCTCTGATGCAA 1583  
Qy 421 AACAGGTTTCCGTCATTATACAAAATTAATCTATTTTAAAGGAAGCAATTAACG 480  
Db 1584 AACAGGTTTCCGTCATTATACAAAATTAATCTATTTTAAAGGAAGCAATTAACG 1643  
Qy 481 ATCCAGTGGTGGCTGCCGTGCCAGTGAATTTCAAGTGTCTCTATCAAGAGGATAGC 540

Db 1644 ATCCAGTGTGGGCTGCTCTCCAAGTGAATTCAGGTGCTCTATCAAGAGGAGATAGC 1703  
Qy 541 AATTACAGTGTGGAGTGTCTTGGCGGCGATGATCTAATATCAAGTCTGATGA 600  
Db 1704 AATTACAGTGTGGAGTGTCTTGGCGGCGATGATCTAATATCAAGTGTGATGA 1763  
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCACAAAGAGCTCCCTTTAGAGCATCACTGTGA 660  
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGCACAAAGAGCTCCCTTTAGAGCATCACTGTGA 1823  
Qy 661 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA 720  
Db 1824 CGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACA 1883  
Qy 721 TTCTTGCTGATCAGTCAGACCTGTGACTTCTTTAATAGTATAGTATAGTAAACAGAGAA 780  
Db 1884 TTCTTGCTGATCAGTCAGACCTGTGACTTCTTTAATAGTATAGTATAGTAAACAGAGAA 1943  
Qy 781 TCACACTGTGTGCTCTTTACAGCTATCAAGTCTCAAGTCTGAGATGACCCAACTTGCAGAAC 840  
Db 1944 TCACACTGTGTGCTCTTTACAGCTATCAAGTCTGAGATGACCCAACTTGCAGAAC 2003  
Qy 841 AAGAGTAATTTGGGCGACCAATTTGGATTTCAGAGTTCAGAGTCTCTCTGACTATATCTCTCC 900  
Db 2004 AAGAGTAATTTGGGCGACCAATTTGGATTTCAGAGTTCAGAGTCTCTCTGACTATATCTCTCC 2063  
Qy 901 AGAAATTTCTCTTTTGAAGTACTACTGGATTGTATGATGGGATGCTCTACAGGCC 960  
Db 2064 AGAAATTTCTCTTTTGAAGTACTACTGGATTGTATGATGGGATGCTCTACAGGCC 2123  
Qy 961 TCATGATCTACAGCCTGGAAAGAAATATCTACTGTCTGTCTATATATGGTGTCTCCTCA 1020  
Db 2124 TCATGATCTACAGCCTGGAAAGAAATATCTACTGTCTGTCTATATATGGTGTCTCCTCA 2179  
Qy 1021 GGTGTGCTATGTGCGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGA 1080  
Db 2180 GGTGTGCTATGTGCGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGA 2239  
Qy 1081 AGTTATATGGGTCACTCTGACAGATGAAAGAGGCTATTTACTTAGATCTGTGGCCAT 1140  
Db 2240 AGTTATATGGGTCACTCTGACAGATGAAAGAGGCTATTTACTTAGATCTGTGGCCAT 2299  
Qy 1141 GCAAGCAGAAAGTTCCCTCTGACCAAAATGCTTTACTGCTCTTACATGATGCTTCTCGGA 1200  
Db 2300 GCAAGCAGAAAGTTCCCTCTGACCAAAATGCTTTACTGCTCTTACATGATGCTTCTCGGA 2359  
Qy 1201 TGAGATGCTCAATTTGACATACAGATATATTTACTGAGTGTCTTTAGTGGGCTGGAAA 1260  
Db 2360 TGAGATGCTCAATTTGACATACAGATATATTTACTGAGTGTCTTTAGTGGGCTGGAAA 2419  
Qy 1261 GCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCCGTAATCGGG 1320  
Db 2420 GCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCCGTAATCGGG 2479  
Qy 1321 AGAACATTTAGAACTGATCTTTTGCACTACCTTCAAGAAAACCTTGATCACTATGTC 1380  
Db 2480 AGAACATTTAGAACTGATCTTTTGCACTACCTTCAAGAAAACCTTGATCACTATGTC 2539  
Qy 1381 TGCTCTAAAGTGAATATATTTTGACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTT 1440  
Db 2540 TGCTCTAAAGTGAATATATTTTGACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTT 2599  
Qy 1441 AACCAATAGGAGGTTTAAATCAACAGAAAACACAGATTTGATCATCATTTTGTATACC 1500  
Db 2600 AACCAATAGGAGGTTTAAATCAACAGAAAACACAGATTTGATCATCATTTTGTATACC 2659  
Qy 1501 TGCCATGTAACTACTCTCTGAAATAAATGCTGTGGTGCATGCGGGTCTTACGGTTTGT 1560  
Db 2660 TGCCATGTAACTACTCTCTGAAATAAATGCTGTGGTGCATGCGGGTCTTACGGTTTGT 2719  
Qy 1561 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATCTCTGAGA 1620  
Db 2720 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATCTCTGAGA 2779

Qy 1621 GACCCAGCAATACCATAGAAATTTACTTAAATAAAAAAAAAAAAAAAAAAAAA 1668  
Db 2780 GACCCAGCAATACCATAGAAATTTACTTAAATAAAAAAAAAAAAAAAAAAAAA 2827  
RESULT 5  
US-09-976-674-20  
; Sequence 20, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akineanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 4676  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-20

Query Match 89.9%; Score 1501; DB 3; Length 4676;  
Best Local Similarity 91.5%; Pred. No. 0;  
Matches 1663; Conservative 0; Mismatches 5; Indels 149; Gaps 1;

Qy 1 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAGATGTCAGAAAATAATGATGCTGA 60  
Db 1164 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAGATGTCAGAAAATAATGATGCTGA 1223  
Qy 61 AGGAAGCATCATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAGATTTCTATTTGA 120  
Db 1224 AGGAAGCATCATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAGATTTCTATTTGA 1283  
Qy 121 AGGAGTTGAATATATTTGCCAGAGCTGGATGGAGCTCCTGAGGGAATAATGCTTGGTCCAT 180  
Db 1284 AGGAGTTGAATATATTTGCCAGAGCTGGATGGAGCTCCTGAGGGAATAATGCTTGGTCCAT 1343  
Qy 181 CCTACTAGATCGCTCCAGAGCTGCCCTACAGATAGTGTGTGATGTCACCTGAATATTTAT 240  
Db 1344 CCTACTAGATCGCTCCAGAGCTGCCCTACAGATAGTGTGTGATGTCACCTGAATATTTAT 1403  
Qy 241 CCCAGTGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATTTCTGT 300  
Db 1404 CCCAGTGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCTGATTTCTGT 1463  
Qy 301 GACCCCACTAATTTCTATGAAGAAACACAGACATCTGGAATAAATATCCATGACATCTT 360  
Db 1464 GACCCCACTAATTTCTATGAAGAAACACAGACATCTGGAATAAATATCCATGACATCTT 1523  
Qy 361 TCATGTTTTTCCCCAAAGTCACGAAGAGAAATGAGTTTATTTTGGCTCTGGAATGCAA 420  
Db 1524 TCATGTTTTTCCCCAAAGTCACGAAGAGAAATGAGTTTATTTTGGCTCTGGAATGCAA 1583  
Qy 421 AACAGGTTTCCGTCATTATATACAAAATTTACATCTATTTTAAAGGAAAGCAATATATAACG 480  
Db 1584 AACAGGTTTCCGTCATTATATACAAAATTTACATCTATTTTAAAGGAAAGCAATATATAACG 1643  
Qy 481 ATCCAGTGGTGGGCTGCTCTCCAGTGTATTTCAAGTGTCTCTATCAAGAGGAGATAGC 540  
Db 1644 ATCCAGTGGTGGGCTGCTCTCCAGTGTATTTCAAGTGTCTCTATCAAGAGGAGATAGC 1703  
Qy 541 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGATCTAATATCCAGTTGATGA 600  
Db 1704 AATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGATCTAATATCCAGTTGATGA 1763

QY 601 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACGTGA 660  
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACGTGA 1823  
QY 661 CGTAGTCAGTACGTAATTCCTGGAGAGGTGACAGAGCTGACTACCGTGGCTACTACA 720  
Db 1824 CGTAGTCAGTACGTAATTCCTGGAGAGGTGACAGAGCTGACTACCGTGGCTACTACA 1883  
QY 721 TTCTTGCTGCATCAGTCAGACCTGCTGCTTTTAAAGTATAGTATACCAAGAGAA 780  
Db 1884 TTCTTGCTGCATCAGTCAGACCTGCTGCTTTTAAAGTATAGTATACCAAGAGAA 1943  
QY 781 TCACACTGTGTCTCCCTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAC 840  
Db 1944 TCACACTGTGTCTCCCTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAC 2003  
QY 841 AAGGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCTCTCTCGACTATATCTCTCC 900  
Db 2004 AAGGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCTCTCTCGACTATATCTCTCC 2063  
QY 901 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCC 960  
Db 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCC 2123  
QY 961 TCATGATCTACAGCTGGAAAGAAATATCTCTACTGTCTGTTCATATATGTTGTC--- 1016  
Db 2124 TCATGATCTACAGCTGGAAAGAAATATCTCTACTGTCTGTTCATATATGTTGTCGGGT 2183  
QY 1017 ----- 1016  
Db 2184 CAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTC 2243  
QY 1017 ----- 1016  
Db 2244 ATTGACTTAGATCGTGTGGGCATCCAGCGCTGGTCTCTATGGAGATACCTCTCCCTGATG 2303  
QY 1017 -----CTCAGGCTGCTATTCCTGCTGGGCCAGCTCATCTCG 1051  
Db 2304 GCATTAAATCGAGGTCAGATATCTTACGGGTTCCTATGCTGGGCCAGCTCATCTCG 2363  
QY 1052 TGGATCTCTATGATACAGGATACACGGAACTTTATATGGTTCACCTGACCAAGATGAA 1111  
Db 2364 TGGATCTCTATGATACAGGATACACGGAACTTTATATGGTTCACCTGACCAAGATGAA 2423  
QY 1112 CAGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGMAACCAAT 1171  
Db 2424 CAGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGMAACCAAT 2483  
QY 1172 CGTTTACTGCTCTTACATGGTTTTCTGGATGAGAAATGTCATTTTGACATACCACTATA 1231  
Db 2484 CGTTTACTGCTCTTACATGGTTTTCTGGATGAGAAATGTCATTTTGACATACCACTATA 2543  
QY 1232 TTACTGAGTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAG 1291  
Db 2544 TTACTGAGTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAG 2603  
QY 1292 AGACAGCATAAAGATTCCTGAATCGGGAGAACATTTATGAATGCTATCTTTGCACTAC 1351  
Db 2604 AGACAGCATAAAGATTCCTGAATCGGGAGAACATTTATGAATGCTATCTTTGCACTAC 2663  
QY 1352 CTTCAAGAAAACCTTTGGATCAGTATTCCTCTCTTAAAGTGATATAATTTTGACCTGTG 1411  
Db 2664 CTTCAAGAAAACCTTTGGATCAGTATTCCTCTCTTAAAGTGATATAATTTTGACCTGTG 2723  
QY 1412 TAGAACTCTCTGATATACACTGGCTATTTTAAACCAATGAGAGGTTTAAATCAACAGAAA 1471  
Db 2724 TAGAACTCTCTGATATACACTGGCTATTTTAAACCAATGAGAGGTTTAAATCAACAGAAA 2783  
QY 1472 CACAGAAATTTGATCACAATTTTGAATACCTGCGCATGTAACATCTACTCTGAAATTAAT 1531  
Db 2784 CACAGAAATTTGATCACAATTTTGAATACCTGCGCATGTAACATCTACTCTGAAATTAAT 2843  
QY 1532 GTGGTGCCATGCAAGGGCTACGGTTTGTGGTAGTATCTAATACCTTAAACCCCATGC 1591

Db 2844 GTGGTGCCATGCAAGGGTCTACGGTTTGTGGTAGTAACTTAATACCTTTAAACCCCATGC 2903  
QY 1592 TCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCAATGAAGATTAATAAAA 1651  
Db 2904 TCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCAATGAAGATTAATAAAA 2963  
QY 1652 AAAAAAAAAAAAAAAAAA 1668  
Db 2964 AAAAAAAAAAAAAAAAAA 2980  
RESULT 6  
US-10-982-512-20  
; Sequence 20, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; PRIORITY FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIORITY FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIORITY FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH 4676  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-982-512-20  
Query Match 89.9%; Score 1501; DB 9; Length 4676;  
Best Local Similarity 91.5%; Pred. No. 0;  
Matches 1663; Conservative 0; Mismatches 5; Indels 149; Gaps 1;  
QY 1 AACAGGTACAGCAAAATCCCTAAAGTGCACCTTTTAAAGATGTCAGAAAATAATGATGATGCTGA 60  
Db 1164 AACAGGTACAGCAAAATCCCTAAAGTGCACCTTTTAAAGATGTCAGAAAATAATGATGATGCTGA 1223  
QY 61 AGGAAGGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATTCCTATTGTA 120  
Db 1224 AGGAAGGATCATAGATGTCATAGATAAGAACTAAATCAACCTTTTGAGATTCCTATTGTA 1283  
QY 121 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGAAAAATATGCTTGGTCCAT 180  
Db 1284 AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGAAAAATATGCTTGGTCCAT 1343  
QY 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTAT 240  
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTAT 1403  
QY 241 CCAGGTAGAAGATGATGTTTATGAAAGGCAGAGACTCAATTCAGTCAGTGCCTGATTCGT 300  
Db 1404 CCAGGTAGAAGATGATGTTTATGAAAGGCAGAGACTCAATTCAGTCAGTGCCTGATTCGT 1463  
QY 301 GAGCCACTAAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360  
Db 1464 GAGCCACTAAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523  
QY 361 TCATGTTTTTCCCCAAAGTCACGAGAGGAATAGGTTTATTTTTCCTCTGATGCAA 420  
Db 1524 TCATGTTTTTCCCCAAAGTCACGAGAGGAATAGGTTTATTTTTCCTCTGATGCAA 1583  
QY 421 AACAGGTTTTCCGTCTATTTATACAAATTTACATCTATTTTAAAGGAGAGCAATATAACG 480  
Db 1584 AACAGGTTTTCCGTCTATTTATACAAATTTACATCTATTTTAAAGGAGAGCAATATAACG 1643

QY 481 ATCCAGTGGTGGGCTGCTCCTCCAAAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGC 540  
Db 1644 ATCCAGTGGTGGGCTGCTCCTCCAAAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGC 1703  
QY 541 AATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCCAAGTTGATGA 600  
Db 1704 AATTACAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCCAAGTTGATGA 1763  
QY 601 AGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACTGTGA 660  
Db 1764 AGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACTGTGA 1823  
QY 661 GGTAGTCAGTACGTAATCCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACA 720  
Db 1824 GGTAGTCAGTACGTAATCCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACA 1883  
QY 721 TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAAGTATAGTACCAAGAGAA 780  
Db 1884 TTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAAGTATAGTACCAAGAGAA 1943  
QY 781 TCCAACACTGTGTGCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCCAAAAC 840  
Db 1944 TCCAACACTGTGTGCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCCAAAAC 2003  
QY 841 AAAGGAATTTTGGGCCACCAATTTTGGCAATTCAGAGGTCTCTTCTGACTATATCTCTCC 900  
Db 2004 AAAGGAATTTTGGGCCACCAATTTTGGCAATTCAGAGGTCTCTTCTGACTATATCTCTCC 2063  
QY 901 AGAAATTTTCTCTTTTGAAGTACTACTGATTTTCAATTTATGATGGATGCTCTACAAGCC 960  
Db 2064 AGAAATTTTCTCTTTTGAAGTACTACTGATTTTCAATTTATGATGGATGCTCTACAAGCC 2123  
QY 961 TCATGATCTACAGCTGGAAAGAAATATCTACTGTCTGTTTTCATATATGGTGTC --- 1016  
Db 2124 TCATGATCTACAGCTGGAAAGAAATATCTACTGTCTGTTTTCATATATGGTGTCGGGT 2183  
QY 1017 ----- 1016  
Db 2184 CAAATAGAAATTCAGCATCAGTGGAAAGGACTCCAATATCTAGCTTCTCGATATGATTTTC 2243  
QY 1017 ----- 1016  
Db 2244 ATTGACTTAGATCGTGTGGGCATCCAGGCTGTCTATGGAGGATACCTCTCCCTGATG 2303  
QY 1017 -----CTCAGGTTGCTATTGCTGGGGCCCCAGTCACTCTG 1051  
Db 2304 GCATTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTG 2363  
QY 1052 TGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTACCCCTGACCCAGCAAGTAA 1111  
Db 2364 TGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTACCCCTGACCCAGCAAGTAA 2423  
QY 1112 CAGGCTATTACTTATGATCTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAAAAT 1171  
Db 2424 CAGGCTATTACTTATGATCTGGCCATGCAAGCAGAAAAGTTCCCTCTGACCAAAAT 2483  
QY 1172 CGTTTACTGCTTTACATGGTTTCTGATGAGAATGTCCATTTTGGACATACCAAGTATA 1231  
Db 2484 CGTTTACTGCTTTACATGGTTTCTGATGAGAATGTCCATTTTGGACATACCAAGTATA 2543  
QY 1232 TTACTGAGTTTTTATGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCCAGGAG 1291  
Db 2544 TTACTGAGTTTTTATGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCCAGGAG 2603  
QY 1292 AGACACAGCATAGAGTTCTGAAATCGGGAGAACATTAAGTGCATCTTTTGCACATAC 1351  
Db 2604 AGACACAGCATAGAGTTCTGAAATCGGGAGAACATTAAGTGCATCTTTTGCACATAC 2663  
QY 1352 CTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAAGTCATATAATTTGACCTGTG 1411  
Db 2664 CTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAAGTCATATAATTTGACCTGTG 2723  
QY 1412 TAGAACTCTCTGTGATACACTGGCTATTTTAAACAAATGAGGAGGTTTAAATCAACAGAAA 1471

Db 2724 TAGAACTCTCTGTGATACACTGGCTATTTTAAACAAATGAGGAGGTTTAAATCAACAGAAA 2783  
QY 1472 CACAGAAATTGATCATCATCATTTTGTATACCTGCGCATGTAACATCTACTCTGAAAAATAAT 1531  
Db 2784 CACAGAAATTGATCATCATCATTTTGTATACCTGCGCATGTAACATCTACTCTGAAAAATAAT 2843  
QY 1532 GTGGTGCCATGCGAGGGTCTACGGTTTGTGGTAGTAAATCTAATACTTAAACCCCATGC 1591  
Db 2844 GTGGTGCCATGCGAGGGTCTACGGTTTGTGGTAGTAAATCTAATACTTAAACCCCATGC 2903  
QY 1592 TCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATACCATAGAAATTACTAAAAA 1651  
Db 2904 TCAAAATCAATGATACATATTCCTGAGAGACCCAGCAATACCATAGAAATTACTAAAAA 2963  
QY 1652 AAAAAAAAAAAAAAAAAA 1668  
Db 2964 AAAAAAAAAAAAAAAAAA 2980

RESULT 7  
US-10-170-789-37  
; Sequence 37, Application US/10170789  
; Publication No. US20030180930A1  
; GENERAL INFORMATION:  
; APPLICANT: Rachel E. Meyers  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Welch, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,  
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF  
; FILE REFERENCE: 1048-191001  
; CURRENT APPLICATION NUMBER: US/10/170, 789  
; CURRENT FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/882,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/212,078  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/934,406  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US01/26052  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,740  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 09/861,801  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: PCT/US01/16549  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 60/205,508  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 09/801,267  
; PRIOR FILING DATE: 2001-03-06  
; PRIOR APPLICATION NUMBER: PCT/US01/07138  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: US 60/187,454  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: US 09/829,671  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: PCT/US01/40483  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: US 60/197,508  
; PRIOR FILING DATE: 2000-04-18  
; PRIOR APPLICATION NUMBER: US 09/961,721

```
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(2874)
US-10-170-789-37

Query Match      81.1%; Score 1353.4; DB 6; Length 3143;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 1; Indels 300; Gaps 1;

Qy      1 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA 60
Db      1179 AACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGA 1238

Qy      61 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCCTATTGGA 120
Db      1239 AGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCCTATTGGA 1298

Qy      121 AGGAGTTGAAATATATTTGCCAGAGCTGGATGGAATCCTCAGGAGAAATATGCTTGGTCCAT 180
Db      1299 AGGAGTTGAAATATATTTGCCAGAGCTGGATGGAATCCTCAGGAGAAATATGCTTGGTCCAT 1358

Qy      181 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAGCTCAATTAATTTAT 240
Db      1359 CCTACTAGATCGCTCCAGACTCGCCTCAGATAGTGTGATCTCAGCTCAATTAATTTAT 1418

Qy      241 CCAAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTGAGTCAGTGCCTGATTCGT 300
Db      1419 CCAAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTGAGTCAGTGCCTGATTCGT 1478

Qy      301 GACGCCACTAATTTATCTATGAGAAACACACAGACATCTGGATAAATATCCATGACATCTT 360
Db      1479 GACGCCACTAATTTATCTATGAGAAACACACAGACATCTGGATAAATATCCATGACATCTT 1538

Qy      361 TCATGTTTTTCCCAAGTCACGAGAGGAATTTGAGTTTATTTTGGCTCTGNAATGCA 420
Db      1539 TCATGTTTTTCCCAAGTCACGAGAGGAATTTGAGTTTATTTTGGCTCTGNAATGCA 1598

Qy      421 AACAGGTTCCGTCATTATATACAAATTCATCTATTTTAAAGGAAGCAATATAACG 480
Db      1599 AACAGGTTCCGTCATTATATACAAATTCATCTATTTTAAAGGAAGCAATATAACG 1658

Qy      481 ATCCAGTGTGGCTGCTCTCCCAAGTGATTTTCAAGTGTCTATCAAAAGAGGATGAC 540
Db      1659 ATCCAGTGTGGCTGCTCTCCCAAGTGATTTTCAAGTGTCTATCAAAAGAGGATGAC 1718

Qy      541 AATTAACAGTGTGAAATTTGGGAAGTTCTTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 600
Db      1719 AATTAACAGTGTGAAATTTGGGAAGTTCTTTGGCCGGCATGGATCTAATATCCAAGTTGATGA 1778

Qy      601 AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCCTTTAGAGCATCACCTGTGA 660
Db      1779 AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCCTTTAGAGCATCACCTGTGA 1838
```

```
Qy      661 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCCGTGGCTACTCACA 720
Db      1839 CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCCGTGGCTACTCACA 1898

Qy      721 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 780
Db      1899 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAACCAAGAGAA 1958

Qy      781 TCCAACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGCAAAAC 840
Db      1959 TCCAACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTCGCAAAAC 2018

Qy      841 AAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCCTCTTCCTGACTATATCTCCTCC 900
Db      2019 AAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCCTCTTCCTGACTATATCTCCTCC 2078

Qy      901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACATTTGTATGGGATGCTCTCAAGGCC 960
Db      2079 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACATTTGTATGGGATGCTCTCAAGGCC 2138

Qy      961 TCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTCTGTTCATATATGTTGGTCTCTCA 1020
Db      2139 TCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTCTGTTCATATATGTTGGTCTCTCA 2198

Qy      1021 ----- 1020
Db      2199 GGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGC 2258

Qy      1021 ----- 1020
Db      2259 CTCTAGGTTATGTGGTTGTAGTAGAACAAGGGGATCCTGTACCAGAGGCTTAA 2318

Qy      1021 ----- 1020
Db      2319 ATTTGAAGCGCCTTTAAATATAAATAATGGGTCAAAATAGAAATTTGACGATCAGGTGGGAAG 2378

Qy      1021 ----- 1020
Db      2379 ACTCAATATCTAGCTTCTCGATATGATTTTCAATTTAGATCGTGTGGGCATCCACGG 2438

Qy      1021 ----- 1020
Db      2439 CTGGTCTCTAGGAGATACCTCTCCCTGATGGCATTAATGACAGAGGTGAGATATCTTCAG 2498

Qy      1021 GGTTCCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1080
Db      2499 GGTTCCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 2558

Qy      1081 ACGTTATATGGGTCACTCCCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCAT 1140
Db      2559 ACGTTATATGGGTCACTCCCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCAT 2618

Qy      1141 GCAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGTCTTTACATGTTTCTCTGGA 1200
Db      2619 GCAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGTCTTTACATGTTTCTCTGGA 2678

Qy      1201 TGAGATGTCATTTTGGCATACAGTATATTTACTGAGTTTCTTTAGTGGGCTGGAAA 1260
Db      2679 TGAGATGTCATTTTGGCATACAGTATATTTACTGAGTTTCTTTAGTGGGCTGGAAA 2738

Qy      1261 GCCATATGATTTACATCTATCTCAGAGAGACACAGCATACAGATTCCTGATCGGG 1320
Db      2739 GCCATATGATTTACATCTATCTCAGAGAGACACAGCATACAGATTCCTGATCGGG 2798

Qy      1321 AGAACATTTAGACTGCACTCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGTTATTC 1380
Db      2799 AGAACATTTAGACTGCACTCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGTTATTC 2858

Qy      1381 TGCTCTAAAAAGTGAATATAAATTTTGAAGCTCTGTAGAACTCTCTGGTATACACTGGCTATTT 1440
Db      2859 TGCTCTAAAAAGTGAATATAAATTTTGAAGCTCTGTAGAACTCTCTGGTATACACTGGCTATTT 2918
```



QY 1441 AACCAATGAGGAGTTAATCAACAGAAACAGAAATGATCATCATCTTTGATACC 1500  
DB 2919 AACCAATGAGGAGTTAATCAACAGAAACAGAAATGATCATCATCTTTGATACC 2978  
QY 1501 TGCCATGTAAACATCTACTCTCTGAAATAAATGTTGGTCCCATGACAGGGGTCTACGGTTGT 1560  
DB 2979 TGCCATGTAAACATCTACTCTCTGAAATAAATGTTGGTCCCATGACAGGGGTCTACGGTTGT 3038  
QY 1561 GGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAATGATACATATTTCTTGAGA 1620  
DB 3039 GGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAATGATACATATTTCTTGAGA 3098  
QY 1621 GACCCGCAATACCAATAGAAATTAATAAAAAAAAAAAAAAAAAAAAA 1665  
DB 3099 GACCCGCAATACCAATAGAAATTAATAAAAAAAAAAAAAAAAAAAAA 3143

RESULT 8  
US-09-976-674-12  
; Sequence 12, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Q1, Scève  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 4829  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-976-674-12

Query Match 80.9%; Score 1349.6; DB 3; Length 4829;  
Best Local Similarity 84.5%; Pred. No. 0;  
Matches 1664; Conservative 0; Mismatches 4; Indels 302; Gaps 1;

QY 1 AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGCTGA 60  
DB 1164 AACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATATGATGCTGA 1223  
QY 61 AGAAGGATCATAGATGTCATAGATAAGAACTAATCAACCTTTTGAGATTCTATTGA 120  
DB 1224 AGAAGGATCATAGATGTCATAGATAAGAACTAATCAACCTTTTGAGATTCTATTGA 1283  
QY 121 AGAAGTGAATATATGTCAGAGCTGGATGCTCTGAGGAAATATGCTGGTCCAT 180  
DB 1284 AGAAGTGAATATATGTCAGAGCTGGATGCTCTGAGGAAATATGCTGGTCCAT 1343  
QY 181 CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCAATATTTAT 240  
DB 1344 CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCAATATTTAT 1403  
QY 241 CCCAGTAGAAGATGATGTTATGGAAGGAGGAGACTCAATGAGTCAGTGGCTGATCTGT 300  
DB 1404 CCCAGTAGAAGATGATGTTATGGAAGGAGGAGACTCAATGAGTCAGTGGCTGATCTGT 1463  
QY 301 GAGGCCATTAATTTCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360  
DB 1464 GAGGCCATTAATTTCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523  
QY 361 TCATGTTTTCCCAAGTACGAGAGAAATTCAGTTATTTTTCCTCTGAATCAA 420  
DB 1524 TCATGTTTTCCCAAGTACGAGAGAAATTCAGTTATTTTTCCTCTGAATCAA 1583  
QY 421 AACAGGTTTCCGTCATTATACAAATTAATCATCTATTTTAAAGGAAACAAATATAACG 480

DB 1584 AACAGGTTTCCGTCATTATATACAAATTAATCATCTATTTTAAAGGAAACAAATATAACG 1643  
QY 481 ATCCAGTGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCCCTATCAAGAGGAGATAGC 540  
DB 1644 ATCCAGTGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCCCTATCAAGAGGAGATAGC 1703  
QY 541 AATTACCAAGTGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCAAGTTGATGA 600  
DB 1704 AATTACCAAGTGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCAAGTTGATGA 1763  
QY 601 AGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACTGTGA 660  
DB 1764 AGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACTGTGA 1823  
QY 661 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACA 720  
DB 1824 CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTACA 1883  
QY 721 TTCTTGTGTCATCAGTCAGCTGCTGACTCTTTTAAAGTAAAGTATAGTAACCAAGAA 780  
DB 1884 TTCTTGTGTCATCAGTCAGCTGCTGACTCTTTTAAAGTAAAGTATAGTAACCAAGAA 1943  
QY 781 TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTTGAAGATGACCAACTTGCAGAAAC 840  
DB 1944 TCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTTGAAGATGACCAACTTGCAGAAAC 2003  
QY 841 AAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGCTCTCTCTGACTATATCTCTCC 900  
DB 2004 AAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGCTCTCTCTGACTATATCTCTCC 2063  
QY 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTATGGATGCTCTCAAGCC 960  
DB 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTATGGATGCTCTCAAGCC 2123  
QY 961 TCATGATCTACAGCTGGAGAAATATCTACTGCTGCTTTCATATATGTTGCTGCTCTCA 1020  
DB 2124 TCATGATCTACAGCTGGAGAAATATCTACTGCTGCTTTCATATATGTTGCTGCTCT 2183  
QY 1021 ----- 1020  
DB 2184 CAGGTGAGTTGGTGAATTAATCGTTTAAAGAGTCAAGTATTTCCGTTGAATACCTTA 2243  
QY 1021 ----- 1020  
DB 2244 GCCTCTCTAGTTATCTGGTTGTAGTATAGACACAGGGATCTCTGTCACGAGGCTT 2303  
QY 1021 ----- 1020  
DB 2304 AAATTTGAAGGCGCTTTAAATATAAATGGGTCAATAGAAATGACGATCAGGTGGAA 2363  
QY 1021 ----- 1020  
DB 2364 GGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGGGATCCAC 2423  
QY 1021 ----- 1020  
DB 2424 GGCTGTCTTATGAGAGGATACCTCTCCCTGATGGCAATTAATGACAGGTGAGATATCTTC 2483  
QY 1021 --GTTCTCTATCTCGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 1078  
DB 2484 AGGTTTCTATCTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 2543  
QY 1079 GAAGCTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCC 1138  
DB 2544 GAAGCTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCC 2603  
QY 1139 ATGCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTCTCTG 1198  
DB 2604 ATGCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTCTCTG 2663  
QY 1199 GATGAGATCTCCATTTTGCACATACAGTATATTAATGAGTTTCTTGTAGGGCTGGA 1258



Db	2664	GATGAGAATGTCCATTTTGCACATACCAGTATATTTACTGAGTCTTTTGTAGTGAGGGCTGGA	2723
Qy	1259	AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAAGATTCCTGAATCG	1318
Db	2724	AAAGCCATATGATTTACAGATCTATCTCAGGAGAGACACAGCATAAAGATTCCTGAATCG	2783
Qy	1319	GGAGAACATATGAACTGCATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGTATT	1378
Db	2784	GGAGAACATATGAACTGCATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGTATT	2843
Qy	1379	CTGCTCTAAAGTGTATTAATTTTGACCTGTGTAGAACTCTCTGGTATACCTGGCTAT	1438
Db	2844	CTGCTCTAAAGTGTATTAATTTTGACCTGTGTAGAACTCTCTGGTATACCTGGCTAT	2903
Qy	1439	TTAAACCAATGAGAGGGTTTAAATCAACAGAAAACACAGAAATGTATCATCATTTTGATA	1498
Db	2904	TTAAACCAATGAGAGGGTTTAAATCAACAGAAAACACAGAAATGTATCATCATTTTGATA	2963
Qy	1499	CTGCGATGTAACTACTCTCTGAAAATAAATGTGTGCCATGCAGGGGTCTACCGTTT	1558
Db	2964	CTGCGATGTAACTACTCTCTGAAAATAAATGTGTGCCATGCAGGGGTCTACCGTTT	3023
Qy	1559	GTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATATATCTCTGA	1618
Db	3024	GTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATATATCTCTGA	3083
Qy	1619	GAGACCCAGCAATACCAATGAATTAATTAACCAATTAATTAACCAATTAATTAACCA	1668
Db	3084	GAGACCCAGCAATACCAATGAATTAATTAACCAATTAATTAACCAATTAATTAACCA	3133
RESULT 9			
US-10-982-512-12			
; Sequence 12, Application US/10982512			
; Publication No. US20050059081A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Khinsanya, Karen			
; APPLICANT: Riviere, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/10/982,512			
; PRIOR FILING DATE: 2004-11-05			
; PRIOR FILING DATE: 2001-10-12			
; PRIOR FILING DATE: 2000-10-11			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 12			
; LENGTH: 4829			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-982-512-12			
Query Match			
Best Local Similarity 80.9%; Score 1349,6; DB 9; Length 4829;			
Matches 1664; Conservative 0; Mismatches 4; Indels 302; Gaps 1;			
Qy	1	AACAGGTACAGCAATCTTAAGTCACCTTTTAAGATGTCAGAAATTAATGATTCATCGTA	60
Db	1164	AACAGGTACAGCAATCTTAAGTCACCTTTTAAGATGTCAGAAATTAATGATTCATCGTA	1223
Qy	61	AGGAGGATCATAGTGTATAGATAGAGAACTAATCAACCTTTGAGATCTATTGTA	120
Db	1224	AGGAGGATCATAGTGTATAGATAGAGAACTAATCAACCTTTGAGATCTATTGTA	1283
Qy	121	AGGAGGATCATAGTGTATAGATAGAGAACTAATCAACCTTTGAGATCTATTGTA	180
Db	1284	AGGAGGATCATAGTGTATAGATAGAGAACTAATCAACCTTTGAGATCTATTGTA	1343
Qy	181	CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATATTATTTAT	240

Db	1344	CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTTAT	1403
Qy	241	CCAGTATAGAAAGATGATGTTATGGAAAGGCAGAGACTCATTTGAGTCAGTGCCTGATTCGT	300
Db	1404	CCAGTATAGAAAGATGATGTTATGGAAAGGCAGAGACTCATTTGAGTCAGTGCCTGATTCGT	1463
Qy	301	GAGCCACTAATTTATCTATGAAAGAACAAACAGACATCTGGATAAATATCCATGACATCTT	360
Db	1464	GAGCCACTAATTTATCTATGAAAGAACAAACAGACATCTGGATAAATATCCATGACATCTT	1523
Qy	361	TCATGTTTTTCCCAAGTCAAGAGAGGAAATTCAGTTTATTTTTTGGCTCTGGAATGCAA	420
Db	1524	TCATGTTTTTCCCAAGTCAAGAGAGGAAATTCAGTTTATTTTTTGGCTCTGGAATGCAA	1583
Qy	421	AACAGGTTTCCGTCATTTATACAAAATTTACATCTAATTTTAAAGGAAGAACAAATTAACG	480
Db	1584	AACAGGTTTCCGTCATTTATACAAAATTTACATCTAATTTTAAAGGAAGAACAAATTAACG	1643
Qy	481	ATCCAGTGTGGGCTGCCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC	540
Db	1644	ATCCAGTGTGGGCTGCCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC	1703
Qy	541	AATTACAGTGTGATGGAATGGGAATTTTGGCGGCGATGGATCTAATATCCAAGTTGATGA	600
Db	1704	AATTACAGTGTGATGGAATGGGAATTTTGGCGGCGATGGATCTAATATCCAAGTTGATGA	1763
Qy	601	AGTCAGAGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATACCTGTGA	660
Db	1764	AGTCAGAGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATACCTGTGA	1823
Qy	661	CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACA	720
Db	1824	CGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACA	1883
Qy	721	TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAAACAGAGAA	780
Db	1884	TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAAACAGAGAA	1943
Qy	781	TCCACACTGTGTCCCTTTTCAAGCTATCAAGTCTCGAAGATGACCACTTTGCAAAAC	840
Db	1944	TCCACACTGTGTCCCTTTTCAAGCTATCAAGTCTCGAAGATGACCACTTTGCAAAAC	2003
Qy	841	AAAGGAATTTTGGGCGCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCC	900
Db	2004	AAAGGAATTTTGGGCGCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCC	2063
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAGCC	960
Db	2064	AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAGCC	2123
Qy	961	TCATGATCTACAGCTGGAAGAAATATCCTGCTGCTTCATATATGTTGGTCTCTCA	1020
Db	2124	TCATGATCTACAGCTGGAAGAAATATCCTGCTGCTTCATATATGTTGGTCTCTCT	2183
Qy	1021	-----	1020
Db	2184	CAGGTGCGAGTTGGTGAATTAATCCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTA	2243
Qy	1021	-----	1020
Db	2244	GCCTCTCTAGGTTATGTTGGTTAGTGAATAGCAACAGGGGATCCTGTACCGAGGGCTT	2303
Qy	1021	-----	1020
Db	2304	AAATTTGAAGGGCCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGTGGAA	2363
Qy	1021	-----	1020
Db	2364	GGACTCCAAATATCTAGCTTCTCGATATGATTTCAITTAGCTGTGGGCGATCCAC	2423
Qy	1021	-----	1020



Db 2304 ATTTGAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAATTCAGCATCAGGTGGAAAG 2363  
Qy 1021 ----- 1020  
Db 2364 ACTCCAATATCTAGCTTCGATATAGATTTTCATTTAGCTTAGATCGTGTGGGCATCCACGG 2423  
Qy 1021 ----- 1020  
Db 2424 CTGGTCCATGAGAGATACCTCTCCCTGATGGCATTAAATGCGAGAGGTGAGATATCTTCAG 2483  
Qy 1021 GGTGTCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1080  
Db 2484 GGTGTCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 2543  
Qy 1081 ACGTATATATGGGTGACCCCTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCAT 1140  
Db 2544 ACGTATATATGGGTGACCCCTGACAGAAATGAAACAGGGCTATTACTTAGGATCTGTGGCCAT 2603  
Qy 1141 GCAAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 1200  
Db 2604 GCAAGCAGAAAAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGA 2663  
Qy 1201 TGAGATGTCCTATTTTGACATACACAGTATATTTACTGAGTCTTTTGTGAGGGCTGGAAA 1260  
Db 2664 TGAGATGTCCTATTTTGACATACACAGTATATTTACTGAGTCTTTTGTGAGGGCTGGAAA 2723  
Qy 1261 GCCATATGATTTACAGATCTATCTCCAGGAGAGACACAGCATTAAGTCTCTGANTCGGG 1320  
Db 2724 GCCATATGATTTACAGATCTATCTCCAGGAGAGACACAGCATTAAGTCTCTGANTCGGG 2783  
Qy 1321 AGACATATGAACTGCATCTTTTGCACTACCTTCAGAGAAACCTTGGATCACTGATTCG 1380  
Db 2784 AGACATATGAACTGCATCTTTTGCACTACCTTCAGAGAAACCTTGGATCACTGATTCG 2843  
Qy 1381 TGCTCTAAAAGTGATATAATTTTGACCTGTGTGAGAACTCTCTGTGATACACTGGCTATTT 1440  
Db 2844 TGCTCTAAAAGTGATATAATTTTGACCTGTGTGAGAACTCTCTGTGATACACTGGCTATTT 2903  
Qy 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATTTTGTATACC 1500  
Db 2904 AACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATCATTTTGTATACC 2963  
Qy 1501 TGCCATGTAACTACTCTCTGAAATAAATGTGTGCGCATGCGGGGTCTACGGTTTGT 1560  
Db 2964 TGCCATGTAACTACTCTCTGAAATAAATGTGTGCGCATGCGGGGTCTACGGTTTGT 3023  
Qy 1561 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 1620  
Db 3024 GGTAGTAATCTAATACCTTAAACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGA 3083  
Qy 1621 GACCAGCAATACCAATGAATTTACTTAAAAA 1657  
Db 3084 GACCAGCAATACCAATGAATTTACTTAAAAA 3120

RESULT 11  
US-10-825-632-2  
; Sequence 2 Application US/10825632  
; Publication N US20040191826A1  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT; Catherine Anne  
; TITLE OF INVENTION: DIFEPIDIL PEPTIDASES  
; FILE REFERENCE: FCSB-100-Div. 1  
; CURRENT APPLICATION NUMBER: US/10/825, 632  
; PRIOR FILING DATE: 2004-04-15  
; PRIOR APPLICATION NUMBER: US 10/070,464  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 3120  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-825-632-2

Query Match 80.7%; Score 1347; DB 8; Length 3120;  
Best Local Similarity 84.7%; Pred. No. 0;  
Matches 1657; Conservative 0; Mismatches 0; Indels 300; Gaps 1;  
Qy 1 AACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGTGTGAGAAATAAATGATGCTGTA 60  
Db 1164 AACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGTGTGAGAAATAAATGATGCTGTA 1223  
Qy 61 AGGAAGGATCATAGATGTATAGATAAGGAACATAATCAACCTTTTGAGATTCATTTTGA 120  
Db 1224 AGGAAGGATCATAGATGTATAGATAAGGAACATAATCAACCTTTTGAGATTCATTTTGA 1283  
Qy 121 AGGAGTTGATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCAT 180  
Db 1284 AGGAGTTGATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCAT 1343  
Qy 181 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAGTCTGATTTAT 240  
Db 1344 CCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAGTCTGATTTAT 1403  
Qy 241 CCNAGTAGAGATGATTTATGGAAGGAGAGACTCATTTGAGTCACTGCTGATTTCTGT 300  
Db 1404 CCNAGTAGAGATGATTTATGGAAGGAGAGACTCATTTGAGTCACTGCTGATTTCTGT 1463  
Qy 301 GAGCCACTAATTTATCTATGAAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360  
Db 1464 GAGCCACTAATTTATCTATGAAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523  
Qy 361 TCATGTTTTTCCCAAGTCAAGAGAGAAATTCAGTTTATTTTGGCTCTGAATGCAA 420  
Db 1524 TCATGTTTTTCCCAAGTCAAGAGAGAAATTCAGTTTATTTTGGCTCTGAATGCAA 1583  
Qy 421 AACAGGTTCCGTCATTTTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATAAAGC 480  
Db 1584 AACAGGTTCCGTCATTTTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATAAAGC 1643  
Qy 481 ATCCAGTGGTGGCTGCCTGCTCCAAGTGTCTTCAAGTGTCTATCAAAAGAGGAGATAGC 540  
Db 1644 ATCCAGTGGTGGCTGCCTGCTCCAAGTGTCTTCAAGTGTCTATCAAAAGAGGAGATAGC 1703  
Qy 541 AATTACCAGTGGTGAATGGAAAGTTCTTGGCCGCGCATGGATCTAATATCCAAGTTGATGA 600  
Db 1704 AATTACCAGTGGTGAATGGAAAGTTCTTGGCCGCGCATGGATCTAATATCCAAGTTGATGA 1763  
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTA 660  
Db 1764 AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTA 1823  
Qy 661 GGTAGTCAGTTACGTAATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACA 720  
Db 1824 GGTAGTCAGTTACGTAATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACA 1883  
Qy 721 TTCTTGTGATCAGTCAGCACTGCTCTTTTAAAGTAAAGTATAGTACACAGAGAA 780  
Db 1884 TTCTTGTGATCAGTCAGCACTGCTCTTTTAAAGTAAAGTATAGTACACAGAGAA 1943  
Qy 781 TCACACTGTGTCTCCCTTTACAGTCTCAAGTCTGAGATGACCCACTTCCAAAC 840  
Db 1944 TCACACTGTGTCTCCCTTTACAGTCTCAAGTCTGAGATGACCCACTTCCAAAC 2003  
Qy 841 AAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCTCC 900  
Db 2004 AAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATCTCTCTCC 2063



Qy 301 GAGCCACTAATTATCTATGAGAACACACAGACATCTGGATAAATATCCATGACATCTT 360  
Db 1453 GAGCCACTAATTATCTATGAGAACACACAGACATCTGGATAAATATCCATGACATCTT 1512  
Qy 361 TCATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTTGGCTCTGAAAGCAA 420  
Db 1513 TCATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTTGGCTCTGAAAGCAA 1572  
Qy 421 AACAGGTTCCGTCATTTATACAAATTAACATCTATTTTAAAGGAAAGCAAAATATAAAG 480  
Db 1573 AACAGGTTCCGTCATTTATACAAATTAACATCTATTTTAAAGGAAAGCAAAATATAAAG 1632  
Qy 481 ATCCAGTGGGGCTGCTGCTCCCAAGTGAATTCAGTGTCTTATCAAAAGAGGAGATAGC 540  
Db 1633 ATCCAGTGGGGCTGCTGCTCCCAAGTGAATTCAGTGTCTTATCAAAAGAGGAGATAGC 1692  
Qy 541 AATACAGTGGTGAATGGAAAGTTCTTGGCCGCGATGGATCTAATATCCAAGTTGATGA 600  
Db 1693 AATACAGTGGTGAATGGAAAGTTCTTGGCCGCGATGGATCTAATATCCAAGTTGATGA 1752  
Qy 601 AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAGCTCCCTTTTAGAGCATCACCTGTA 660  
Db 1753 AGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAGCTCCCTTTTAGAGCATCACCTGTA 1812  
Qy 661 CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 720  
Db 1813 CGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA 1872  
Qy 721 TCTCTGCTGATCAGTCAGCAGCTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAGAA 780  
Db 1873 TCTCTGCTGATCAGTCAGCAGCTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAGAA 1932  
Qy 781 TCCACACTGTGTCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTTGCAAAAC 840  
Db 1933 TCCACACTGTGTCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTTGCAAAAC 1992  
Qy 841 ABAGGAATTTGGGCCACCATTTTGGATTTCAGAGGCTCTCTCTGACTATACCTCCG 900  
Db 1993 ABAGGAATTTGGGCCACCATTTTGGATTTCAGAGGCTCTCTCTGACTATACCTCCG 2052  
Qy 901 AGAAATTTCTCTTTTGAAGTACTCTGAAATTTTACATGTATGAGGATGCTCTACAGGC 960  
Db 2053 AGAAATTTCTCTTTTGAAGTACTCTGAAATTTTACATGTATGAGGATGCTCTACAGGC 2112  
Qy 961 TCATGATCTACAGCTCGGAAAGAAATATCCTACTGTCTGTCTATATATGAGGCTCTCA 1020  
Db 2113 TCATGATCTACAGCTCGGAAAGAAATATCCTACTGTCTGTCTATATATGAGGCTCTCA 2172  
Qy 1021 ----- 1020  
Db 2173 GGTGCAAGTGGTAATAATCGGTTTAAAGGAGTCAAGTATTTTCGGTTGAATACCCTAGC 2232  
Qy 1021 ----- 1020  
Db 2233 CTCCTAGGTTATGTGTTGTAGTAGACAAACAGGGGATCCTGTCCACGAGGGCTTAA 2292  
Qy 1021 ----- 1020  
Db 2293 ATTTGAAGGCGCTTTAAATATAAAATGGGTCAAATAGAAATGACGATCAGGTGAAGG 2352  
Qy 1021 ----- 1020  
Db 2353 ACTCCAATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACG 2412  
Qy 1021 ----- 1020  
Db 2413 CTGGTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTTCAGATATCTTCAG 2472  
Qy 1021 GGTGCTATTTGGGGCCCCAGTCACCTCTGTGATCTCTTATGATACAGGATACACGGA 1080  
Db 2473 GGTGCTATTTGGGGCCCCAGTCACCTCTGTGATCTCTTATGATACAGGATACACGGA 2532

Qy 1081 ACCTTATATGGGTCAACCTGACCAAGATGACAGGGCTATTACTTAGGATCTGTGGCCAT 1140  
Db 2533 ACCTTATATGGGTCAACCTGACCAAGATGACAGGGCTATTACTTAGGATCTGTGGCCAT 2592  
Qy 1141 GCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTTACATGATGTTTCCCTGGA 1200  
Db 2593 GCAAGCAGAAAAAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTTACATGATGTTTCCCTGGA 2652  
Qy 1201 TGAGAAATGTCATTTTGGCAATACCAAGTATATTTACTGAGTTTTTTAGTGAGGGCTGGAAA 1260  
Db 2653 TGAGAAATGTCATTTTGGCAATACCAAGTATATTTACTGAGTTTTTTAGTGAGGGCTGGAAA 2712  
Qy 1261 GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATAAAGAGTTCCTGAATCGGG 1320  
Db 2713 GCCATATGATTTACAGATCTATCCTCAGGAGAGACACAGCATAAAGAGTTCCTGAATCGGG 2772  
Qy 1321 AGAACATTTAGAACTGCATCTTTTGCACATACCTTCAAGAAAAACCTTGGATCACGTTATTC 1380  
Db 2773 AGAACATTTAGAACTGCATCTTTTGCACATACCTTCAAGAAAAACCTTGGATCACGTTATTC 2832  
Qy 1381 TGCTCTAAAAGTGATATAAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 1440  
Db 2833 TGCTCTAAAAGTGATATAAATTTTGGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTT 2892  
Qy 1441 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTCATCATCATTTTGTATAC 1500  
Db 2893 AACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATTCATCATCATTTTGTATAC 2952  
Qy 1501 TGCCATGTAACATCTACTCTCTGAAATAAATGTGGTGCATGACAGGGCTCTACGTTTGT 1560  
Db 2953 TGCCATGTAACATCTACTCTCTGAAATAAATGTGGTGCATGACAGGGCTCTACGTTTGT 3012  
Qy 1561 GGTAGTAATCTAATCAATTAACCCACATGCTCAAAATCAATGATCATATTTCTCTGAGA 1620  
Db 3013 GGTAGTAATCTAATCAATTAACCCACATGCTCAAAATCAATGATCATATTTCTCTGAGA 3072  
Qy 1621 GACCCAGCAATACCATAGAAATTTACTTAAAAAA 1654  
Db 3073 GACCCAGCAATACCATAGAAATTTACTTAAAAAA 3106

RESULT 13  
US-09-976-674-22  
; Sequence 22, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 4685  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-22

Query Match 80.5%; Score 1343.6; DB 3; Length 4685;  
Best Local Similarity 87.0%; Pred. No. 0;  
Matches 1589; Conservative 0; Mismatches 79; Indels 158; Gaps 4;

Qy 1 AACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAAGTCTCAGAATATATGATGATCTGA 60  
Db 1164 AACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAAGTCTCAGAATATATGATGATCTGA 1223  
Qy 61 AGGAAGATCATAGATGTCATAGATAGGAATAATCAACCTTTTGAGATTTCTATTGA 120

1224 AGAAGGATCAAGATGTCATAGATAAGGAACAAATCAACCTTTTGGAGATCTATTGCA 1283  
121 AGAGTTGAAATATATGCGAGCTGGATGCACTCTGAGGGAATATGCTTGGTCCAT 180  
1284 AGGAGTTGAATATATGCGAGCTGGATGCACTCTGAGGGAATATGCTTGGTCCAT 1343  
181 CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATTAATTTAT 240  
1344 CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATTAATTTAT 1403  
241 CCAGTAGAAGATGATGTTATGGAAGGACAGACTCAATGAGTCAAGTGCCTGATCTGT 300  
1404 CCAGTAGAAGATGATGTTATGGAAGGACAGACTCAATGAGTCAAGTGCCTGATCTGT 1463  
301 GAGGCCACTAAATTAATCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360  
1464 GAGGCCACTAAATTAATCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1523  
361 TCATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTGGCTCTGAATGCAA 420  
1524 TCATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTGGCTCTGAATGCAA 1583  
421 AACAGTTTCCGTCATTTATATCAAAATTAATCTATTTTAAAGGAAGCAATATATAACG 480  
1584 AACAGTTTCCGTCATTTATATCAAAATTAATCTATTTTAAAGGAAGCAATATATAACG 1643  
481 ATCCAGTGTGGGCTGCTGCTCCAGTGAATTCAGTGTCCCTATCAAGAGAGATAGC 540  
1644 ATCCAGTGTGGGCTGCTGCTCCAGTGAATTCAGTGTCCCTATCAAGAGAGATAGC 1703  
541 AATTACAGTGTGATGGAAGTCTTTGGCGGCGATGATCTAATATCCAAGTTCATGA 600  
1704 AATTACAGTGTGATGGAAGTCTTTGGCGGCGATGATCTAATATCCAAGTTCATGA 1763  
601 AGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACTGTA 660  
1764 AGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACTGTA 1823  
661 CGTAGTCAGTTAGTTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA 720  
1824 CGTAGTCAGTTAGTTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTACA 1883  
721 TTCTGCTGATCAGTCAGTCAGCTGTGACTCTTTTAAAGTAAATAGTAAACCAAGAGAA 780  
1884 TTCTGCTGATCAGTCAGTCAGCTGTGACTCTTTTAAAGTAAATAGTAAACCAAGAGAA 1943  
781 TCACACATGTGTGTCCTTTTCAAGCTATCAAGTCTGAGATGACCCAACTTTGCAAAAC 840  
1944 TCACACATGTGTGTCCTTTTCAAGCTATCAAGTCTGAGATGACCCAACTTTGCAAAAC 2003  
841 AAAGGAATTTGGGCGACCACTTTTGGATTTCAG---CAGTCTCTCTCTGACTATATCTC 896  
2004 AAAGGAATTTGGGCGACCACTTTTGGATTTCAGTCTCTCAGTGCAGTGGTGAATATCGG 2063  
897 CTCA-----GAAATTTTCTCTTTTGAAGTACTGAGTAAATTTACATTTGATGGGATGC 950  
2064 TTTTAAAGGAGTCAAGTAATTTCCGCTTGAATACCCCTAGCTCTTAGGTTAAGTGTGTA 2123  
951 TCTACAGCCTCATGATCTACAG-----GAGTCTCTCTCTGACTATATCTC----- 973  
2124 GTGATACAAACAGGGGATCTGTCCAGGGGCTTAAATTTGAAGGCGCTTTAAATAT 2183  
974 -----GAGTCTCTCTCTGACTATATCTC-----GAGTCTCTCTCTGACTATATCTC 973  
2184 AAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGTCTCGA 2243  
974 -----CCTGGAAGAAATATCTCTACTGTCTGTTTATATATGTTGGTCTCTC 1019  
2244 TATGATTTTCAATGACTTAGATCGTGTGGGATCCAGGCTGGTCTTATGGAGGATCTC 2303  
1020 -----AGGTTGCTATTTGCTGGGCGCCCA 1042

Db 2304 TCCTGATGGCAATTAATGACAGAGTCAAGATATCTTCAGGGTTCCTATTGCTGGGGCCCCA 2363  
Qy 1043 GTCACCTCTGTGATCTTTCTATGATACAGATAACAGGAACGTTTATATGGGTCAACCTGAC 1102  
Db 2364 GTCACCTCTGTGATCTTTCTATGATACAGATAACAGGAACGTTTATATGGGTCAACCTGAC 2423  
Qy 1103 CAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCT 1162  
Db 2424 CAGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCT 2483  
Qy 1163 GAACCAAAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGCACAT 1222  
Db 2484 GAACCAAAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGCACAT 2543  
Qy 1223 ACCAGTATATTAAGTGTGTTTACTGAGGCTGGAAAGCCATATGATTTTACAGATCTAT 1282  
Db 2544 ACCAGTATATTAAGTGTGTTTACTGAGGCTGGAAAGCCATATGATTTTACAGATCTAT 2603  
Qy 1283 CCTCAGGAGACACAGCATTAAGATTCCTGAATCGGGAGAACATTAATGACATGCTCTT 1342  
Db 2604 CCTCAGGAGACACAGCATTAAGATTCCTGAATCGGGAGAACATTAATGACATGCTCTT 2663  
Qy 1343 TTGCACCTACTCTTCAAGAAAACCTTCGATACGTTTGTCTCTTAAAAAGTGATATAATTT 1402  
Db 2664 TTGCACCTACTCTTCAAGAAAACCTTCGATACGTTTGTCTCTTAAAAAGTGATATAATTT 2723  
Qy 1403 TGACCTGTGTAGAACTCTCTGATATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATC 1462  
Db 2724 TGACCTGTGTAGAACTCTCTGATATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATC 2783  
Qy 1463 AACAGAAAACACAGAAATTCATCATCACTTTTGTGATACCTGCCATGTAACATCTACTCTG 1522  
Db 2784 AACAGAAAACACAGAAATTCATCATCACTTTTGTGATACCTGCCATGTAACATCTACTCTG 2843  
Qy 1523 AAAATAAATGTGTGTCATGTCAGGGCTCTTACGTTTGTGTAGTAAATCTAATACCTTAAAC 1582  
Db 2844 AAAATAAATGTGTGTCATGTCAGGGCTCTTACGTTTGTGTAGTAAATCTAATACCTTAAAC 2903  
Qy 1583 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATAGAAT 1642  
Db 2904 CCCACATGCTCAAAATCAAAATGATACATATTTCTGAGAGACCCAGCAATACCATAGAAT 2963  
Qy 1643 TACTAAAAAATAAAAAAAAAAAAAA 1668  
Db 2964 TACTAAAAAATAAAAAAAAAAAAAA 2989

RESULT 14

US-10-982-512-22  
; Sequence 22, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; CURRENT FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 4685  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-982-512-22

Query Match 80.5%; Score 1343.6; DB 9; Length 4685;



Best Local Similarity 87.08; Pred. No. 0; Matches 1589; Conservative 0; Mismatches 79; Indels 158; Gaps 4;									
QY	1	AACAGGTACAGCAAAATCCTAAAGTCATCTTTAAGATGTCAGAAATAATGATTTGATCTGA	60						
DB	1164	AACAGGTACAGCAAAATCCTAAAGTCATCTTTAAGATGTCAGAAATAATGATTTGATCTGA	1223						
QY	61	AGGAAGGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAGATTTCTATTGA	120						
DB	1224	AGGAAGGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAGATTTCTATTGA	1283						
QY	121	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTGAGGGGAAAAATATGCTTGGTCCAT	180						
DB	1284	AGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTGAGGGGAAAAATATGCTTGGTCCAT	1343						
QY	181	CCTACTAGATTCGGTCCCGACAGCTCGCTACAGATAGTGTGATCTCCACCTGAATATTAT	240						
DB	1344	CCTACTAGATTCGGTCCCGACAGCTCGCTACAGATAGTGTGATCTCCACCTGAATATTAT	1403						
QY	241	CCCAAGTAAAGATGATGTTATGGAAGGCAGAGACTCAATTCAGTCAGTCGCTGATTCGT	300						
DB	1404	CCCAAGTAAAGATGATGTTATGGAAGGCAGAGACTCAATTCAGTCAGTCGCTGATTCGT	1463						
QY	301	GAGCCACTAATTTATCTATGAGNACACAGACATCTGGATAATATCCATGACATCTT	360						
DB	1464	GAGCCACTAATTTATCTATGAGNACACAGACATCTGGATAATATCCATGACATCTT	1523						
QY	361	TCATGTTTTTCCCAAGTCACGAGAGAAATGAGTTATTTTTTGGCTCTGAATGCAA	420						
DB	1524	TCATGTTTTTCCCAAGTCACGAGAGAAATGAGTTATTTTTTGGCTCTGAATGCAA	1583						
QY	421	AACAGGTTTCGGTCAATTATACAAATTAATCAATTTTAAAGGAAACCAATAATAACG	480						
DB	1584	AACAGGTTTCGGTCAATTATACAAATTAATCAATTTTAAAGGAAACCAATAATAACG	1643						
QY	481	ATCCAGTGTGGGCTGCCTGCTCCAAAGTATTTCAAAGTGCTTATCAAGAGGAGATAGC	540						
DB	1644	ATCCAGTGTGGGCTGCCTGCTCCAAAGTATTTCAAAGTGCTTATCAAGAGGAGATAGC	1703						
QY	541	AATTACAGTGTGTAATGGAAAGTCTTTGGCCGGCATGGATCTAATATCCAAGTTGATGA	600						
DB	1704	AATTACAGTGTGTAATGGAAAGTCTTTGGCCGGCATGGATCTAATATCCAAGTTGATGA	1763						
QY	601	AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCCTTTAGAGCATCACCTGTA	660						
DB	1764	AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCCTTTAGAGCATCACCTGTA	1823						
QY	661	CSTAGTCAGTTACGTAATCTCTGGAGAGGTGCACAAGGCTGACTGACCGTGGCTACTACA	720						
DB	1824	CSTAGTCAGTTACGTAATCTCTGGAGAGGTGCACAAGGCTGACTGACCGTGGCTACTACA	1883						
QY	721	TTCTTGCTGCATCAGTCAGCATGTGACTCTTTTAAGTAAGTATAGTAACCAAGAA	780						
DB	1884	TTCTTGCTGCATCAGTCAGCATGTGACTCTTTTAAGTAAGTATAGTAACCAAGAA	1943						
QY	781	TCCACACTGTGTGCCCTTTACAGCTATCAAGTCCTGAGATGACCCCACTTGCAAAAC	840						
DB	1944	TCCACACTGTGTGCCCTTTACAGCTATCAAGTCCTGAGATGACCCCACTTGCAAAAC	2003						
QY	841	AAAGGAATTTGGGCCACATTTTGGATTGAG-----CAGGTCTCTTCCCTGACTATACTC	896						
DB	2004	AAAGGAATTTGGGCCACATTTTGGATTGAG-----CAGGTCTCTTCCCTGACTATACTC	2063						
QY	897	CTCCA-----GAAATTTTCTTTGGAAGTACTACTGGATTTACATGTTATGGGATGC	950						
DB	2064	TTTAAAGGAGTCAGATTTTCCGCTTGAATACCCCTAGCTCTCTAGGTTATGTGGTTGA	2123						
QY	951	TCTACAGCTCATGATCTACAG-----	973						
DB	2124	GTGATAGACACAGGGGATCTGTCTACCGAGGGCTTAAATTTTGAAGGCGCTTTAAATAT	2183						
QY	974	-----	973						



FILE REFERENCE: PI-0085 USN  
CURRENT APPLICATION NUMBER: US/10/275,505  
CURRENT FILING DATE: 2002-11-04  
PRIOR APPLICATION NUMBER: PCT/US01/14651  
PRIOR FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: 60/209,402  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: 60/207,477  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: 60/205,803  
PRIOR FILING DATE: 2000-05-17  
PRIOR APPLICATION NUMBER: 60/203,566  
PRIOR FILING DATE: 2000-05-11  
PRIOR APPLICATION NUMBER: 60/202,082  
PRIOR FILING DATE: 2000-05-04  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: PERL Program  
SEQ ID NO 16  
LENGTH: 2510  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No: 376067CB1  
US-10-275-505-16

Query Match: 74.9%; Score 1250.6; DB 7; Length 2510;  
Best Local Similarity 87.6%; Pred. No. 0;  
Matches 1482; Conservative 0; Mismatches 44; Indels 165; Gaps 4;

Qy	1	AACAGGTACAGCAAAATCTTAAGTCACTTTTAAAGATGTACAGAAATATGATGATCTGA	60
Db	831	AACAGGTACAGCAAAATCTTAAGTCACTTTTAAAGATGTACAGAAATATGATGATCTGA	890
Qy	61	AGGAAGGATCATAGATGTATAGATAAGGAACATAATCAACCTTTTGAATTCATTTGA	120
Db	891	AGGAAGGATCATAGATGTATAGATAAGGAACATAATCAACCTTTTGAATTCATTTGA	950
Qy	121	AGGAGTGAATATATTTGACAGCTGGATGAGCTCCTGAGGAAATATGCTGGTCCAT	180
Db	951	AGGAGTGAATATATTTGACAGCTGGATGAGCTCCTGAGGAAATATGCTGGTCCAT	1010
Qy	181	CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT	240
Db	1011	CTTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTAT	1070
Qy	241	CCAGTAGAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATTCGT	300
Db	1071	CCAGTAGAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCTGATTCGT	1130
Qy	301	GAGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCT	360
Db	1131	GAGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCT	1190
Qy	361	TCATGTTTTTCCCAAGTCAAGAGAGAAATTCAGTTTATTTTGGCTCTGAATGCAA	420
Db	1191	TCATGTTTTTCCCAAGTCAAGAGAGAAATTCAGTTTATTTTGGCTCTGAATGCAA	1250
Qy	421	AACAGGTTCCGTCATTTATACAAATTCATCTATTTTAAAGGAAGCAATAAAG	480
Db	1251	AACAGGTTCCGTCATTTATACAAATTCATCTATTTTAAAGGAAGCAATAAAG	1310
Qy	481	ATCCAGTGGTGGCTGGCTCCAGTCAATTTCAAGTGTCTATCAAGAGGAGATAGC	540
Db	1311	ATCCAGTGGTGGCTGGCTCCAGTCAATTTCAAGTGTCTATCAAGAGGAGATAGC	1369
Qy	541	AATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGATGGATCTAATATCCAAAGTTGATGA	600
Db	1370	TAGGAAC-----TCCATCCTGTATGTGTGTGACATATATGTTGATGCCAAGTTGATGA	1424
Qy	601	AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTA	660
Db	1425	AGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTA	1484

Qy	661	CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA	720
Db	1485	CGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACA	1544
Qy	721	TTCTTGTCTGCATCAGTCAGCAGCTGTGACCTTTTATAAGTAAGTATAGTAAACCAAGAA	780
Db	1545	TTCTTGTCTGCATCAGTCAGCAGCTGTGACCTTTTATAAGTAAGTATAGTAAACCAAGAA	1604
Qy	781	TCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAGATGACCCAACTTGCAGAAAC	840
Db	1605	TCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAGATGACCCAACTTGCAGAAAC	1664
Qy	841	AAAGGAATTTTGGGCCACCACTTTTGAATTCAGACAGTCTCTTCTGCTACTATATCTCTCC	900
Db	1665	AAAGGAATTTTGGGCCACCACTTTTGAATTCAGACAGTCTCTTCTGCTACTATATCTCTCC	1724
Qy	901	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAAGCC	960
Db	1725	AGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAAGCC	1784
Qy	961	TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATGTTGGTCTCTCA	1020
Db	1785	TCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGCTGTTTCATATATGTTGGTCTCTCA	1844
Qy	1021	-----	1020
Db	1845	GGTGCAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTTCCGCTTGAATACCCTAGC	1904
Qy	1021	-----	1020
Db	1905	CTCTCTAGGTTATGTGTTGTAGTATGATACAAACAGGGGATCCTGTACCGAGGGCTTAA	1964
Qy	1021	-----GGTTCCTATTTGCTGGGGCCCGAGTCACCTCTGTG	1053
Db	1965	ATTTGAAGGCGCTTTAAATATATAAATATGTTGCTGCTGGGGCCCGAGTCACCTCTGTG	2024
Qy	1054	GATCTTCTATGATACAGGATACACGGAAGTTATATGGGTCAACCTGACCAAGATGAACA	1113
Db	2025	GATCTTCTATGATACAGGATACACGGAAGTTATATGGGTCAACCTGACCAAGATGAACA	2084
Qy	1114	GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCG	1173
Db	2085	GGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCG	2144
Qy	1174	TTTACTGCTCTTATCATGTTTCTGATGAGAAATGTCATTTTGGACATATACAGATATAT	1233
Db	2145	TTTACTGCTCTTATCATGTTTCTGATGAGAAATGTCATTTTGGACATATACAGATATAT	2204
Qy	1234	ACTGAGTTTTTATGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAG	1293
Db	2205	ACTGAGTTTTTATGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAG	2252
Qy	1294	ACAGCATATAGAGTTCCCTGAATCGGAGAAACATTTATGAACTGATCTTTTGGCACTACT	1353
Db	2253	ACAGCATATAGAGTTCCCTGAATCGGAGAAACATTTATGAACTGATCTTTTGGCACTACT	2312
Qy	1354	TCAAGAAAACCTTGGATCACGATATGCTCTCTTAAAGTATATATTTTGAACCTGTGTA	1413
Db	2313	TCAAGAAAACCTTGGATCACGATATGCTCTCTTAAAGTATATATTTTGAACCTGTGTA	2372
Qy	1414	GAACTCTCTGGTATACATGCTGCTATTTTAAACAAATGAGGAGTTTAAATCAACAGAAAA	1473
Db	2373	GAACTCTCTGGTATACATGCTGCTATTTTAAACAAATGAGGAGTTTAAATCAACAGAAAA	2432
Qy	1474	CAGAAATGATCATCATATTTTGTATACCTGCAATGTAAATCTCTCTGAAATAAATGT	1533
Db	2433	CAGAAATGATCATCATATTTTGTATACCTGCAATGTAAATCTCTCTGAAATAAATGT	2492
Qy	1534	GGTGCCATGCA	1544
Db	2493	GGTGCCATGAA	2503

Search completed: May 4, 2006, 11:21:52  
Job time : 1496.73 secs

-----

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:49:39 ; Search time 2279.71 Seconds  
(without alignments)  
2981.544 Million cell updates/sec

Title: us-10-825-632-6  
Perfect score: 1669  
Sequence: 1 aaacggtacagcaaatccta.....aaaaaaaaaaaaaaaaaa 1669

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA\_New:  
1: /SIDSS/ptodata/2/pubpna/US08\_NEW\_PUB.seq1:  
2: /SIDSS/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
3: /SIDSS/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
4: /SIDSS/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
5: /SIDSS/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
6: /SIDSS/ptodata/2/pubpna/US03\_NEW\_PUB.seq:  
7: /SIDSS/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
8: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
9: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
10: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq1:  
11: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq2:  
12: /SIDSS/ptodata/2/pubpna/US10\_NEW\_PUB.seq3:  
13: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
14: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
15: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq2:  
16: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq3:  
17: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq4:  
18: /SIDSS/ptodata/2/pubpna/US11\_NEW\_PUB.seq5:  
19: /SIDSS/ptodata/2/pubpna/US60\_NEW\_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1353.4	81.1	3143	17	US-11-151-601-19
2	1087.4	65.2	2649	17	US-11-151-601-21
C 3	118	7.1	1346	7	US-09-925-065A-669313
C 4	113.6	6.8	609	7	US-09-925-065A-743558
C 5	113.6	6.8	609	7	US-09-925-065A-743559
C 6	73.4	4.4	612	7	US-09-925-065A-818446
7	51.6	3.1	2238	18	US-11-079-463-1186
8	50.6	3.0	378	11	US-10-932-182A-81332
9	50.6	3.0	378	11	US-10-932-182A-81332
10	49.6	3.0	2778	11	US-10-932-182A-5649
11	49.6	3.0	2778	11	US-10-932-182A-5649
12	45.6	2.7	3332	18	US-11-208-288-1
13	45.6	2.7	3407	11	US-10-501-035-34
14	44.8	2.7	1624	10	US-10-131-826A-181
					Sequence 19, Appl
					Sequence 21, Appl
					Sequence 669313,
					Sequence 743558,
					Sequence 743559,
					Sequence 818446,
					Sequence 1186, Ap
					Sequence 81332, A
					Sequence 81332, A
					Sequence 5649, Ap
					Sequence 5649, Ap
					Sequence 1, Appli
					Sequence 34, Appl
					Sequence 181, Appl

15	44.8	2.7	1624	11	US-10-973-115B-181	Sequence 181, App
16	44.8	2.7	1624	13	US-10-137-873A-181	Sequence 181, App
17	44.8	2.7	1624	13	US-10-152-370-181	Sequence 181, App
18	44.8	2.7	1624	18	US-11-290-153-181	Sequence 181, App
19	44	2.6	2217	18	US-11-208-288-3	Sequence 3, Appli
20	44	2.6	2301	10	US-10-522-789-1	Sequence 1, Appli
21	42	2.5	2283	18	US-11-208-288-5	Sequence 5, Appli
22	41	2.5	588	7	US-09-925-065A-486279	Sequence 486279, A
23	41	2.5	622	7	US-09-925-065A-49906	Sequence 49906, A
24	41	2.5	622	11	US-10-301-480-151144	Sequence 151144, A
25	41	2.5	622	12	US-10-301-480-764553	Sequence 764553, A
26	40.4	2.4	600	17	US-11-136-527-6535	Sequence 6535, Ap
27	40.4	2.4	1825	17	US-11-136-527-2439	Sequence 2439, Ap
28	40	2.4	660	13	US-10-194-487-467	Sequence 467, App
29	40	2.4	660	13	US-10-195-883-467	Sequence 467, App
30	40	2.4	660	13	US-10-195-888-467	Sequence 467, App
31	40	2.4	660	13	US-10-195-889-467	Sequence 15, Appl
32	40	2.4	660	13	US-10-226-486-15	Sequence 179, App
33	40	2.4	666	13	US-10-986-405-179	Sequence 143, App
34	40	2.4	670	13	US-10-986-405-143	Sequence 143, App
35	39.8	2.4	600	10	US-10-750-185-3494	Sequence 3494, Ap
36	39.8	2.4	600	10	US-10-750-623-3494	Sequence 3494, Ap
37	39.8	2.4	4852	17	US-11-136-527-2130	Sequence 2130, Ap
C 38	39.8	2.4	5152	10	US-10-240-708-47	Sequence 47, Appl
C 39	39.8	2.4	170995	17	US-11-121-086-35	Sequence 35, Appl
40	39.6	2.4	1400	17	US-11-136-527-6859	Sequence 6859, Ap
41	39.6	2.4	3708	17	US-11-136-527-2763	Sequence 2763, Ap
42	39.4	2.4	555	7	US-09-925-065A-394749	Sequence 394749, A
43	39.4	2.4	1516	9	US-10-505-928-579	Sequence 579, App
44	39.4	2.4	1944	13	US-10-782-413-61	Sequence 61, Appl
C 45	39	2.3	1819	10	US-10-750-185-38461	Sequence 38461, A

ALIGNMENTS

RESULT 1  
US-11-151-601-19  
; Sequence 19, Application US/11151601  
; Publication No. US20060003413A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Weich, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE  
; FILE REFERENCE: AND PROTEASE FAMILY MEMBERS AND USES THEREOF  
; FILE REFERENCE: MP100-054P1RCPIOMNIDVIM  
; CURRENT APPLICATION NUMBER: US/11/151,601  
; CURRENT FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: US 10/170,789  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/892,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/212,078  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/934,406  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US01/26052  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,740  
; PRIOR FILING DATE: 2000-08-21

; Remaining Prior Application data removed - See File Wrapper or PALM.

```

: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 19
: LENGTH: 3143
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (229)...(2874)
US-11-151-601-19

```

Query Match 81.1%; Score 1353.4; DB 17; Length 3143;  
Best Local Similarity 84.7%; Pred. No. 9.8e-289;  
Matches 1664; Conservative 0; Mismatches 1; Indels 300; Gaps 1;

1	QY	1	AACGGTACAGCAAAATCCTTAAAGTCTACTTTTAAAGATGTGAGAAATAATGATTGATGCTGA	60
1179	DB	1179	AACAGGTACAGCAAAATCCTTAAAGTCTACTTTTAAAGATGTGAGAAATAATGATTGATGCTGA	1238
61	QY	61	AGGAAGGATCATAGATCTCATAGATAAGGAACTAAATTCACCTTTTGAGATTCCTATTGGA	120
1239	DB	1239	AGGAAGGATCATAGATGTCTAGATAAGGAACTAAATTCACCTTTTGAGATTCCTATTGGA	1298
1321	QY	1321	AGGAGTTGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAAAATAATGCTTGGTCCAT	180
1299	DB	1299	AGGAGTTGAAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAAAATAATGCTTGGTCCAT	1358
181	QY	181	CCTACTAGATCGCTCCAGAGACTCGCTCACAGATAGTGTGATCTCACCTCGAATATTATTAT	240
1359	DB	1359	CCTACTAGATCGCTCCAGAGACTCGCTCCAGATAGTGTGATCTCACCTCGAATATTATTAT	1418
241	QY	241	CCCAGTAGAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCCTGATTCGT	300
1419	DB	1419	CCCAGTAGAGATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGCCCTGATTCGT	1478
301	QY	301	GACGCCCTAAATTATCTATGAAGAAACAACAGACATCTGGATAAAATATCCAATGACATCTT	360
1479	DB	1479	GACGCCCTAAATTATCTATGAAGAAACAACAGACATCTGGATAAAATATCCAATGACATCTT	1538
361	QY	361	TCATGTTTTTCCCACAAAGTCCAGGAGGAAAATGAGTTTATTTTGGCTCTGAATCGAA	420
1539	DB	1539	TCATGTTTTTCCCACAAAGTCCAGGAGGAAAATGAGTTTATTTTGGCTCTGAATCGAA	1598
421	QY	421	AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATAAACAAG	480
1599	DB	1599	AACAGGTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATAAACAAG	1658
481	QY	481	ATCCAGTGGTGGGCTGCTCTCCAAAGTGATTTCAAAGTGCTCTATCAAAGAGGAGATAGC	540
1659	DB	1659	ATCCAGTGGTGGGCTGCTCTCCAAAGTGATTTCAAAGTGCTCTATCAAAGAGGAGATAGC	1718
541	QY	541	AATTAACAGTGGTGAATGGGAAGTCTCTGGCGGCATGGATCTAATATCCAAGTTGATGA	600
1719	DB	1719	AATTAACAGTGGTGAATGGGAAGTCTCTGGCGGCATGGATCTAATATCCAAGTTGATGA	1778
601	QY	601	AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTGAAGCATCACCTGTA	660
1779	DB	1779	AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTGAAGCATCACCTGTA	1838
661	QY	661	CGTAGTCAGTATGCTAAATCTGGAGAGGTGACAAGGCTGATGACCGTGGCTACTACACA	720
1839	DB	1839	CGTAGTCAGTATGCTAAATCTGGAGAGGTGACAAGGCTGATGACCGTGGCTACTACACA	1898
721	QY	721	TTCTTGTGTCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAGAA	780
1899	DB	1899	TTCTTGTGTCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAGAA	1958
781	QY	781	TCCACACTGTGTCCCTTTTACAAGCTATCAAGTCCCTGAAGTACCCCAACTTTGCAAAAC	840
1959	DB	1959	TCCACACTGTGTCCCTTTTACAAGCTATCAAGTCCCTGAAGTACCCCAACTTTGCAAAAC	2018
841	QY	841	AAAGGAATTTTGGGCCACCAATTTTGGATTTCAAGCAGGTCCTCTCTGACTATATCTCTCC	900

Db 3099 GA<sup>CC</sup>CAGCAATACCATRAGAATTACTAAAAA<sup>AAAAAAAAAAAAAAAAAAAA</sup> 3143

## RESULT 2

US-11-151-601-21

; Sequence 21, Application US/11151601

; Publication No. US2006003413A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Meyers, Rachel E.

; APPLICANT: Olandt, Peter J.

; APPLICANT: Kapeller-Libermann, Rosana

; APPLICANT: Curtis, Rory A. J.

; APPLICANT: Williamson, Mark

; APPLICANT: Welch, Nadine

; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,

; FILE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF

; FILE REFERENCE: MP100-054P1RCP10MIDVIM

; CURRENT APPLICATION NUMBER: US/11/151,601

; CURRENT FILING DATE: 2005-06-13

; PRIOR APPLICATION NUMBER: US 10/170,789

; PRIOR FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: US 09/797,039

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: PCT/US01/06525

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US 60/186,061

; PRIOR FILING DATE: 2000-02-29

; PRIOR APPLICATION NUMBER: US 09/882,166

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: PCT/US01/19269

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 60/212,078

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: US 09/934,406

; PRIOR FILING DATE: 2001-08-21

; PRIOR APPLICATION NUMBER: PCT/US01/26052

; PRIOR FILING DATE: 2001-08-21

; PRIOR APPLICATION NUMBER: US 60/226,740

; PRIOR FILING DATE: 2000-08-21

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 45

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 21

; LENGTH: 2649

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-151-601-21

Query Match 65.2%; Score 1087.4; DB 17; Length 2649;  
Best Local Similarity 82.3%; Pred. No. 5e-230;  
Matches 1398; Conservative 0; Mismatches 1; Indels 300; Gaps 1;

Qy 1 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATGATGCTGA 60

Db 951 AACAGGTACAGCAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATGATGCTGA 1010

Qy 61 AGGAGGATCATAGATGTCATAGATAGGAACCTAAATCAACCTTTTGAGATTCATTATGA 120

Db 1011 AGGAGGATCATAGATGTCATAGATAGGAACCTAAATCAACCTTTTGAGATTCATTATGA 1070

Qy 121 AGGAGTGAATATATGTCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGCTCCAT 180

Db 1071 AGGAGTGAATATATGTCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGCTCCAT 1130

Qy 181 CCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCAACCTGAATATTTAT 240

Db 1131 CCTACTAGATCGCTCCAGACTCGCCTGAGATAGTGTGATCTCAACCTGAATATTTAT 1190

Qy 241 CCCAGTAGAAGATGATGTTATGAAAGGCGAGACCTCATTTGAGTCAAGTGCCTGATTCGT 300

Db 1191 CCCAGTAGAAGATGATGTTATGAAAGGCGAGACCTCATTTGAGTCAAGTGCCTGATTCGT 1250

Qy 301 GACGCCACTAATTAATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 360

Db 1251 GACGCCACTAATTAATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTT 1310

Qy 361 TCATGTTTTTCCCAAGTCAAGAGGAAATGAGTTTATTTTAAAGGAAGCAATATATAACG 480

Db 1311 TCATGTTTTTCCCAAGTCAAGAGGAAATGAGTTTATTTTAAAGGAAGCAATATATAACG 1370

Qy 421 AACAGGTTTCGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATATAACG 480

Db 1371 AACAGGTTTCGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAATATATAACG 1430

Qy 481 ATCCAGTGTGGGCTGCTCTCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC 540

Db 1431 ATCCAGTGTGGGCTGCTCTCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGC 1490

Qy 541 AATTACCAAGTGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 600

Db 1491 AATTACCAAGTGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 1550

Qy 601 AGTCAGAAAGGCTGGTATATTTTGAAGSCACCAAGACTCCCTTTAGAGCATCACTGTA 660

Db 1551 AGTCAGAAAGGCTGGTATATTTTGAAGSCACCAAGACTCCCTTTAGAGCATCACTGTA 1610

Qy 661 GGTAGTCAGTTACGTAATCTCTGAGAGGTGACAAAGGCTGACGACCGTGGCTACTCACA 720

Db 1611 GGTAGTCAGTTACGTAATCTCTGAGAGGTGACAAAGGCTGACGACCGTGGCTACTCACA 1670

Qy 721 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAA 780

Db 1671 TTCTTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAA 1730

Qy 781 TCCACACTGTGTGTCCTTTTACAAGCTATCAAGCTCTCTGAAGATGACCCAACTTGCAAAAC 840

Db 1731 TCCACACTGTGTGTCCTTTTACAAGCTATCAAGCTCTCTGAAGATGACCCAACTTGCAAAAC 1790

Qy 841 AAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCTGACTATACCTCTCC 900

Db 1791 AAAGGAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTCTGACTATACCTCTCC 1850

Qy 901 AGAAATTTTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCC 960

Db 1851 AGAAATTTTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCC 1910

Qy 961 TCATGATCTACAGCTCGGAAAGAAATATCTACTGCTGTTTCATATATATGTTGGTCTCTCA 1020

Db 1911 TCATGATCTACAGCTCGGAAAGAAATATCTACTGCTGTTTCATATATATGTTGGTCTCTCA 1970

Qy 1021 ----- 1020

Db 1971 GGTGAGTTGGTGAATAATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGC 2030

Qy 1021 ----- 1020

Db 2031 CTCTCTAGGTTATGTGTTGTGATGATAGCAACAGGGGATCCTGTCTACCCGAGGCTTAA 2090

Qy 1021 ----- 1020

Db 2091 ATTTGAAGCGCCTTTAAATATAAATGGGTCAATAGAAATTGACGATCAGGTGGAAGG 2150

Qy 1021 ----- 1020

Db 2151 ACTCCAAATATCTAGCTTCTCGATATGATTTCATTGACTTAGATCGTGTGGGCATCCACGG 2210

Qy 1021 ----- 1020

Db 2211 CTGGTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTGAGATATCTTCAG 2270

Qy 1021 ----- 1020

Db 2271 GGTGCTATTCCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACCGA 2330

Qy 1081 AGTTATATGGGTCAACCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGCCAT 1140

```
Db 2331 ACCTTATATGGGTCACCCCTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGCCAT 2390
Qy 1141 GCAAGCAGAAAAGTTCCTCTCGAACCAATCGTTTACTGCTCTTACATGCTTTCTCTGGA 1200
Db 2391 GCAAGCAGAAAAGTTCCTCTCGAACCAATCGTTTACTGCTCTTACATGCTTTCTCTGGA 2450
Qy 1201 TGAGAAATGTCATTTTGCATACACAGATATATTACTGAGTTTCTTTAGTGAGGGCTGAAA 1260
Db 2451 TGAGAAATGTCATTTTGCATACACAGATATATTACTGAGTTTCTTTAGTGAGGGCTGAAA 2510
Qy 1261 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCTGAATCGGG 1320
Db 2511 GCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCTGAATCGGG 2570
Qy 1321 AGACATATTAGTAAGTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGCTATTGC 1380
Db 2571 AGACATATTAGTAAGTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGCTATTGC 2630
Qy 1381 TGCTCTAAAAGTCATATA 1399
Db 2631 TGCTCTAAAAGTCATATA 2649
```

```
RESULT 3
US-09-925-065A-669313/C
; Sequence 669313, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669313
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-669313
```

```
Query Match 7.1%; Score 118; DB 7; Length 1346;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 873 CAGTCCTCTTCTGACTACTACTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGAT 932
Db 118 CAGTCCTCTTCTGACTACTACTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGAT 99
Qy 933 TTACATGTTATGGGATGCTCTACAGCCTCATGATCTACAGCTCGAAGAAATATCC 990
Db 58 TTACATGTTATGGGATGCTCTACAGCCTCATGATCTACAGCTCGAAGAAATATCC 1
```

```
RESULT 4
US-09-925-065A-743558/C
; Sequence 743558, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
```

```
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 743558
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743558

Query Match 6.8%; Score 113.6; DB 7; Length 609;
Best Local Similarity 96.7%; Pred. No. 3.2e-15;
Matches 116; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1159 CTCGTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGAATGTCCATTTTGC 1218
Db 205 CTCGAGACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGAATGTCCATTTTGC 146
Qy 1219 ACATACCAGTATATTACTGAGTTTCTTGTAGGGCTCGAAAGCCATATGATTACAGAT 1278
Db 145 ACATACCAGTATATTACTGAGTTTCTTGTAGGGCTCGAAAGCCATATGATTACAGAT 86
```

```
RESULT 5
US-09-925-065A-743559/C
; Sequence 743559, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 743559
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-743559
```

```
Query Match 6.8%; Score 113.6; DB 7; Length 609;
Best Local Similarity 96.7%; Pred. No. 3.2e-15;
Matches 116; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1159 CTCGTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGAATGTCCATTTTGC 1218
Db 205 CTCGAGACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGAATGTCCATTTTGC 146
Qy 1219 ACATACCAGTATATTACTGAGTTTCTTGTAGGGCTCGAAAGCCATATGATTACAGAT 1278
```

Db	145	ACATACAGTATATTACTGAGTTTCTAGTGAGGGCTGGAAGCCATGATGATTACAGGT	86
RESULT 6			
US-09-925-065A-818446/c			
; Sequence 818446, Application US/09925065A			
; Publication No. US20040181048A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, David G.			
; TITLE OF INVENTION: Identification and Mapping of Single			
; Nucleotide Polymorphisms in the Human Genome			
; FILE REFERENCE: 108827.135			
; CURRENT APPLICATION NUMBER: US/09/925,065A			
; CURRENT FILING DATE: 2001-08-08			
; PRIOR APPLICATION NUMBER: US 60/243,096			
; PRIOR FILING DATE: 2000-10-24			
; PRIOR APPLICATION NUMBER: US 60/252,147			
; PRIOR FILING DATE: 2000-11-20			
; PRIOR APPLICATION NUMBER: US 60/250,092			
; PRIOR FILING DATE: 2000-11-30			
; PRIOR APPLICATION NUMBER: US 60/261,766			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/289,846			
; PRIOR FILING DATE: 2001-05-09			
; NUMBER OF SEQ ID NOS: 957086			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 818446			
; LENGTH: 612			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-925-065A-818446			
Query Match			
Best Local Similarity 4.4%; Score 73.4; DB 7; Length 612;			
Matches 74; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1204	GAATGTCATTTTGCACATACAGTATATTACTGAGTTTCTAGTGAGGGCTGGAAGCC	1263
Db	612	GAATGTCATTTTGCACATACAGTATATTACTGAGTTTCTAGTGAGGGCTGGAAGCC	553
Qy	1264	ATATGATTACAGAT	1278
Db	552	ATATGATTACAGGT	538
RESULT 7			
US-11-079-463-1186			
; Sequence 1186, Application US/11079463			
; Publication No. US20060073161A1			
; GENERAL INFORMATION:			
; APPLICANT: Gary L. Breton			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR			
; FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: PAT00-03DIV2			
; CURRENT APPLICATION NUMBER: US/11/079,463			
; CURRENT FILING DATE: 2005-03-14			
; PRIOR APPLICATION NUMBER: US 60/128,705			
; PRIOR FILING DATE: 1999-04-09			
; PRIOR APPLICATION NUMBER: US 09/540,209			
; PRIOR FILING DATE: 2000-04-04			
; NUMBER OF SEQ ID NOS: 10444			
; SEQ ID NO 1186			
; LENGTH: 2238			
; TYPE: DNA			
; ORGANISM: B.fragilis			
US-11-079-463-1186			
Query Match			
Best Local Similarity 3.1%; Score 51.6; DB 18; Length 2238;			
Matches 166; Conservative 0; Mismatches 159; Indels 9; Gaps 1;			
Qy	1029	TTGCTGGGGCCCCGACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTTATA	1088

Db	1907	TTGCTGTGGCGCCACCTACAGACTGGAAATATTACGATACAGTATATACCGAAGCTTTA	1966
Qy	1089	TGGGTCACTCCCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAG	1148
Db	1967	TGCGCAGCGCGAAGAAATGCCGAGGCTATTAAGCAGCTTCAGACTTCAGCCGTGCG	2026
Qy	1149	AAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATG	1208
Db	2027	ACAACTGTCATGGTAAACC-----TGCTCTTGTACACGTTATGGCAGATGATAATG	2077
Qy	1209	TCCATTTTGACATACCACTGATATATTACTGAGTTTTTTAGTGAGGGCTGGAAGCCATATG	1268
Db	2078	TTCACTTCCAGAACTGTACAGAAATATGCAGAGCACTGTGTAACAATCGGAAAAACAGTTTCG	2137
Qy	1269	ATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAGAGTTCTCTGAATCGGGAGAACATT	1328
Db	2138	ATATGAGGTTATACACCAACCGGAATCATAGATCTATGTTGGAATACCCGTAACCACT	2197
Qy	1329	ATGAACCTGCATCTTTTGGCACTACCTTCAAGAAAA	1362
Db	2198	TGTATACGAAGCTGACGAACCTCTTCCGGAATAA	2231
RESULT 8			
US-10-932-182A-81332			
; Sequence 81332, Application US/10932182A			
; Publication No. US20060046253A1			
; GENERAL INFORMATION:			
; APPLICANT: NAKAO, YOSHIHIRO			
; APPLICANT: NAKAMURA, NORIHISA			
; APPLICANT: KODAMA, YUKIKO			
; APPLICANT: FUJIMURA, TOMOKO			
; APPLICANT: ASHIKARI, TOSHIHIKO			
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS			
; FILE REFERENCE: 030685-043			
; CURRENT APPLICATION NUMBER: US/10/932,182A			
; CURRENT FILING DATE: 2004-09-02			
; NUMBER OF SEQ ID NOS: 197023			
; SOFTWARE: PatentIn version 3.3			
; SEQ ID NO 81332			
; LENGTH: 378			
; TYPE: DNA			
; ORGANISM: Saccharomyces pastorianus			
US-10-932-182A-81332			
Query Match			
Best Local Similarity 3.0%; Score 50.6; DB 11; Length 378;			
Matches 117; Conservative 0; Mismatches 89; Indels 3; Gaps 1;			
Qy	1028	ATTGCTGGGGCCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTAT	1087
Db	1	ATGCTCTCGGCGCAGTAACTGACATTTGATGATTCGTTTACACGGAAGATAT	60
Qy	1088	ATGGGTCACTCCCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCA	1147
Db	61	ATGAACCAACCATCGGAAAAACCATGAAGCTATTTTGAAGTATCCA---CCATCCAAAT	117
Qy	1148	GAAGAATTCCTCTGNAACCAATCTTTACTGCTTTACATGTTTCTCGATCGAGAT	1207
Db	118	TTCAAAATCATTCGAAATCTTTTAAAGCGTCTTTTCATTTGTGACGGAACCTTTGATGATAAT	177
Qy	1208	GTCCATTTTGCACATACCACTGATATATTACT	1236
Db	178	GTTCAATACAAATACTTTTAGATTAGT	206
RESULT 9			
US-10-932-182A-81332			
; Sequence 81332, Application US/10932182A			
; Publication No. US20060046253A1			
; GENERAL INFORMATION:			
; APPLICANT: NAKAO, YOSHIHIRO			



; APPLICANT: NAKAMURA, NORIHIRO  
; APPLICANT: KODAMA, YUKIO  
; APPLICANT: FUJIMURA, TOMOKO  
; APPLICANT: ASHIKARI, TOSHIHIKO  
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
; FILE REFERENCE: 030685-043  
; CURRENT APPLICATION NUMBER: US/10/932,182A  
; CURRENT FILING DATE: 2004-09-02  
; NUMBER OF SEQ ID NOS: 197023  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 81332  
; LENGTH: 378  
; TYPE: DNA  
; ORGANISM: Saccharomyces pastorianus  
US-10-932-182A-81332

Query Match 3.0%; Score 50.6; DB 11; Length 378;  
Best Local Similarity 56.8%; Pred. No. 0.23;  
Matches 117; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

Qy 1028 ATTGCTGGGCCCCAGTCTCTGTGGATCTTTCTATGATACAGGATACAGGAACGTTAT 1087  
Db 1 ATGGCTGTGCGCCAGTAAACAACTGGACATTGTATGTCCTGTTTACCGGAAAGATAT 60  
Qy 1088 ATGGCTCACCTGACCAGAAATGAACAGGGCTATTACTTATAGGATCTGTGGCCATGCAAGCA 1147  
Db 61 ATGAACCAACCATCGGAAACCAATGAAGGCTATTTTGAAGTATCCA--CCATCCAAAT 117  
Qy 1148 GAAAGATTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGCTTTCTGGATGAGAAAT 1207  
Db 118 TTCAATCATCTGAAATCTTTAAAGCGTCTTTTCATTGTGACGGAACCTTTCGATGATAAT 177  
Qy 1208 GTCCATTTTGACATACCAAGTATATTACT 1236  
Db 178 GTTCACATACAAATCTTTTAGATTAGT 206

RESULT 10  
US-10-932-182A-5649  
; Sequence 5649, Application US/10932182A  
; Publication No. US20060046253A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAO, YOSHIHIRO  
; APPLICANT: NAKAMURA, NORIHIRO  
; APPLICANT: KODAMA, YUKIO  
; APPLICANT: FUJIMURA, TOMOKO  
; APPLICANT: ASHIKARI, TOSHIHIKO  
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
; FILE REFERENCE: 030685-043  
; CURRENT APPLICATION NUMBER: US/10/932,182A  
; CURRENT FILING DATE: 2004-09-02  
; NUMBER OF SEQ ID NOS: 197023  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 5649  
; LENGTH: 2778  
; TYPE: DNA  
; ORGANISM: Saccharomyces pastorianus  
US-10-932-182A-5649

Query Match 3.0%; Score 49.6; DB 11; Length 2778;  
Best Local Similarity 55.8%; Pred. No. 0.65;  
Matches 116; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

Qy 1026 CTATTGCTGGGCCCCAGTCTCTGTGGATCTTTCTATGATACAGGATACAGGAACGTT 1085  
Db 2414 CAATGGCTGTGACCACTGAACTGGAATCGCTGAAAGCGACTCTTTGTTATGCACGGGACCTCTCTGATGATA 2473  
Qy 1086 ATATGGTCCACCTGACCAAGATGAACAGGGCTATTAGGATCTGTGGCCATGCAAG 1145  
Db 2474 ATATGAATCAACCATCGGAAATGAAGGGTTATTTTGACATATCTA---CTATTAAA 2530  
Qy 1146 CAGAAAGTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGATGAGA 1205

Db 2531 ATTACAAAGTCTTTCGAAATCGTGAAGCGACTCTTTGTTATGCACGGGACCTCTGATGATA 2590  
Qy 1206 ATGTCCTATTTCACATACACAGTATATT 1233  
Db 2591 ATGTTCAATATCAAAATACGTTTAGATT 2618

RESULT 11  
US-10-932-182A-5649  
; Sequence 5649, Application US/10932182A  
; Publication No. US20060046253A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAO, YOSHIHIRO  
; APPLICANT: NAKAMURA, NORIHIRO  
; APPLICANT: KODAMA, YUKIO  
; APPLICANT: FUJIMURA, TOMOKO  
; APPLICANT: ASHIKARI, TOSHIHIKO  
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
; FILE REFERENCE: 030685-043  
; CURRENT APPLICATION NUMBER: US/10/932,182A  
; CURRENT FILING DATE: 2004-09-02  
; NUMBER OF SEQ ID NOS: 197023  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 5649  
; LENGTH: 2778  
; TYPE: DNA  
; ORGANISM: Saccharomyces pastorianus  
US-10-932-182A-5649

Query Match 3.0%; Score 49.6; DB 11; Length 2778;  
Best Local Similarity 55.8%; Pred. No. 0.65;  
Matches 116; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

Qy 1026 CTATTGCTGGGCCCCAGTCTCTGTGGATCTTTCTATGATACAGGATACAGGAACGTT 1085  
Db 2414 CAATGGCTGTGACCACTGAACTGGAATCGCTGAAAGCGACTCTTTGTTATGCACGGGACCTCTCTGATGATA 2473  
Qy 1086 ATATGGTCCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAG 1145  
Db 2474 ATATGAATCAACCATCGGAAATGAAGGGTTATTTTGACATATCTA---CTATTAAA 2530  
Qy 1146 CAGAAAGTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGATGAGA 1205  
Db 2531 ATTACAAAGTCTTTCGAAATCGCTGAAAGCGACTCTTTGTTATGCACGGGACCTCTCTGATGATA 2590  
Qy 1206 ATGTCCTATTTCACATACACAGTATATT 1233  
Db 2591 ATGTTCAATATCAAAATACGTTTAGATT 2618

RESULT 12  
US-11-208-288-1  
; Sequence 1, Application US/11208288  
; Publication No. US20060051366A1  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, Chiwen  
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF  
; FILE REFERENCE: 39533-0001  
; CURRENT APPLICATION NUMBER: US/11/208,288  
; CURRENT FILING DATE: 2005-08-18  
; PRIOR APPLICATION NUMBER: US 60/605,013  
; PRIOR FILING DATE: 2004-08-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH 3332  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-11-208-288-1

Query Match 2.7%; Score 45.6; DB 18; Length 3332;  
Best Local Similarity 46.1%; Pred. No. 5.2;

Matches 188; Conservative 0; Mismatches 219; Indels 1; Gaps 1;

QY 1020 AGTTTGTCTATGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGG 1079

Db 1943 AGTGTGGAAATAGCCGTGGCGCTGTATCCGGTGGGAGTACTATGACTCAGTGTACACAG 2002

QY 1080 AACGTTATATGGTGCACCCCTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGGGCA 1139

Db 2003 AACGTTATATGGTCTCCCACTCCAGAGCAACCTTGACCAATTACAGAAATTCACACAG 2062

QY 1140 TCGAAGCAGAAAGTTCCCTCTCTGAACCAATGTTTACTGTCTTACATGATGTTTCCCTGG 1199

Db 2063 TCATGAGCAGAGCTGAAAATTTAAACAAGTTGAGTACCTCTTATTCATGGAACAGCAG 2122

QY 1200 ATGAGAATGTCATTTTGCATACATACAGTATATCTAGTGTATTTTGTAGGAGGCTGGAA 1259

Db 2123 ATGATAACGTTTCACCTTCAGCAGTCAGTCCAGATCTCCAAAGCCCTGGTGGTGTGGAG 2182

QY 1260 AGCCATATGTTTACAGATCTATCTCAGGAGACACAGACATAGAGTTCCTGAAATCGG 1319

Db 2183 TGGATTCCAGCAATGTTGATATCTATGATGAGACCATGGAATAGTAGCAGCAGCAGC 2242

QY 1320 GAGAACATTATGAACATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCAAGTATTG 1379

Db 2243 ACCACATATATATACCCACATGAGCCACTTCATAAAACAATGTTTCTCTTTACCT-TAG 2301

QY 1380 CTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTAT 1427

Db 2302 CACCTCAAAATACCATGCAATTTAAAGCTTATTAAACTCATTTTGT 2349

RESULT 13

US-10-501-035-34

; Sequence 34, Application US/10501035

; Publication No. US20060046249A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING

; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE

; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS

; FILE REFERENCE: D0185 PCT

; CURRENT APPLICATION NUMBER: US/10/501,035

; PRIOR FILING DATE: 2004-07-09

; PRIOR APPLICATION NUMBER: US 60/350,061

; PRIOR FILING DATE: 2002-01-18

; NUMBER OF SEQ ID NOS: 795

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 34

; LENGTH: 3407

; TYPE: DNA

; ORGANISM: Homo sapiens

Query Match 2.7%; Score 45.6; DB 11; Length 3407;

Best Local Similarity 46.1%; Pred. No. 5.3;

Matches 188; Conservative 0; Mismatches 219; Indels 1; Gaps 1;

QY 1020 AGTTTGTCTATGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGG 1079

Db 2018 AGTGTGGAATAGCGTGGCGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACACAG 2077

QY 1080 AACGTTATATGGTGCACCCCTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCA 1139

Db 2078 AACGTTATATGGTCTCCCACTCCAGAGCAACCTTGACCAATTACAGAAATTCACACAG 2137

QY 1140 TCGAAGCAGAAAGTTCCCTCTCTGAACCAATGTTTACTGTCTTTACATGTTTCTCTGG 1199

Db 2138 TCATGAGCAGAGCTGAAAATTTTAAACAAGTTGAGTACCTCTTATTCATGGAACAGCAG 2197

QY 1200 ATGAGAATGTCATTTTGCATACATACAGTATATTAAGTGTATTTTGTAGTGGGCTGGAA 1259

Db 2198 ATGATAACGTTTCACCTTTCAGCAGTCAGTCCAGATCTCCAAAGCCCTGGTGCATGTGGAG 2257

QY 1260 AGCCATATGATTATACAGATCTATCTCAGGAGACACAGACATAGAGTTCCTGAAATCGG 1319

Db 2258 TGGATTCCAGGCAATGTTGATATCTGATGAAGACCATGGAATAGCTAGCAGCAGCAGC 2317

QY 1320 GAGAACATTATGAACATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCAAGTATTG 1379

Db 2318 ACCACATATATATACCCACATGAGCCACTTCATAAAACAATGTTTCTCTTTACCT-TAG 2376

QY 1380 CTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGGTAT 1427

Db 2377 CACCTCAAAATACCATGCAATTTAAAGCTTATTAAACTCATTTTGT 2424

RESULT 14

US-10-131-826A-181

; Sequence 181, Application US/10131826A

; Publication No. US20050245730A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tamas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330RJC128

; CURRENT APPLICATION NUMBER: US/10/131,826A

; CURRENT FILING DATE: 2002-04-24

; PRIOR APPLICATION NUMBER: 60/049911

; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: 60/056974

; PRIOR FILING DATE: 1997-08-26

; PRIOR APPLICATION NUMBER: 60/059113

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059115

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059117

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059122

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059184

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059352

; PRIOR FILING DATE: 1997-09-19

; PRIOR APPLICATION NUMBER: 60/059588

; PRIOR FILING DATE: 1997-09-19

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 181

; LENGTH: 1624

; TYPE: DNA

; ORGANISM: Homo Sapien

; FEATURE:

; NAME/KEY: unsure

; LOCATION: 1560-1561

; OTHER INFORMATION: unknown base

US-10-131-826A-181

Query Match 2.7%; Score 44.8; DB 10; Length 1624;

Best Local Similarity 63.2%; Pred. No. 6.5;  
Matches 67; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
QY 1564 AGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATGATACATATTCCTGAGAC 1623  
DB 1514 AGAAATTAATGTTGAATCCAAATCCTAGAAATGAATGAGTTTNNTTCCAAAAAAA 1573  
QY 1624 CCAGCAATACCAATGAAGATTACTTAAAAAATCAATGATACATATTCCTGAGAC 1669  
DB 1574 AAA 1619

RESULT 15  
US-10-973-115B-181  
; Sequence 181, Application US/10973115B  
; Publication No. US20060040351A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Desrogers, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING  
; FILE REFERENCE: 39870-3330R1C300C1  
; CURRENT APPLICATION NUMBER: US/10/973,115B  
; CURRENT FILING DATE: 2004-10-22  
; PRIOR APPLICATION NUMBER: US 10/145,747  
; PRIOR FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: US 10/028,072  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: PCT/US00/32678  
; PRIOR FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: US 09/581,742  
; PRIOR FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: PCT/US00/05746  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 60/135,736  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: US 60/123,090  
; PRIOR FILING DATE: 1999-03-05  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 181  
; LENGTH: 1624  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 1560-1561  
; OTHER INFORMATION: unknown base  
US-10-973-115B-181

Query Match 2.7%; Score 44.8; DB 11; Length 1624;  
Best Local Similarity 63.2%; Pred. No. 6.5;  
Matches 67; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
QY 1564 AGTAATCTAATACCTTAAACCCACATGCTCAAAATCAATGATACATATTCCTGAGAC 1623  
DB 1514 AGAAATTAATGTTGAATCCAAATCCTAGAAATGAATGAGTTTNNTTCCAAAAAAA 1573  
QY 1624 CCAGCAATACCAATGAAGATTACTTAAAAAATCAATGATACATATTCCTGAGAC 1669

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 00:41:48 ; Search time 50.6744 Seconds  
(without alignments)  
2894.257 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 2957  
Sequence: 1 aacaggtacagcaatccta.....aaaaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abss/ABSSWEB\_spool/US10825632/runat\_01052006\_105944\_3214/app\_query.fasta\_1  
-DB=A Geneseq -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HCST=abs02p  
-USER=US10825632 @CGN\_1\_1\_605 @runat\_01052006\_105944\_3214 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2482	83.9	465	4 AAB47189	AAB47189 Human DPP
2	2482	83.9	724	5 ABB97362	ABB97362 Novel hum
3	2482	83.9	782	5 ABB97361	ABB97361 Novel hum
4	2447.5	82.8	831	6 ABU92026	ABU92026 Human pro
5	2441.5	82.6	632	4 AAB93565	AAB93565 Human pro
6	2431.5	82.2	738	8 ADT04044	ADT04044 Human pro
7	2422	81.9	824	6 ABU92030	ABU92030 Human pro
8	2422	81.9	882	4 AAB47187	AAB47187 Human DPP
9	2422	81.9	882	5 ABG61591	ABG61591 Human DPP

10	2422	81.9	882	5 AAG78415	AAG78415 Amino aci
11	2422	81.9	882	5 AAE24170	AAE24170 Human dip
12	2422	81.9	882	5 AAU74749	AAU74749 Human pro
13	2422	81.9	882	5 ADI17086	ADI17086 Human NOV
14	2422	81.9	882	6 ABU07720	ABU07720 Human ser
15	2411	81.5	587	5 ADR41398	ADR41398 Human CD-
16	2326	78.7	883	5 ADI17085	ADI17085 Murine NO
17	2244.5	75.9	580	5 AAE14337	AAE14337 Human pro
18	2134.5	72.2	516	6 ABU92029	ABU92029 Human pro
19	1992	67.4	746	8 ADI16334	ADI16334 Human pro
20	1957.5	66.2	690	5 ABG61594	ABG61594 Human DPR
21	1915	64.8	493	7 ADE78977	ADE78977 Human pro
22	1808	61.1	658	5 ABG61600	ABG61600 Human DPR
23	1808	61.1	661	5 ABG61596	ABG61596 Human DPR
24	1540	52.1	613	5 ABG61601	ABG61601 Human DPR
25	1494.5	50.5	689	5 ADI17084	ADI17084 Human NOV
26	1494.5	50.5	755	8 ADQ67811	ADQ67811 Novel hum
27	1494.5	50.5	830	5 AAE24171	AAE24171 Human dip
28	1494.5	50.5	863	5 ABG61592	ABG61592 Human DPP
29	1494.5	50.5	863	5 ADI17083	ADI17083 Human NOV
30	1494.5	50.5	892	5 ABG61602	ABG61602 Human DPR
31	1494.5	50.5	892	5 ABG61604	ABG61604 Human DPR
32	1494.5	50.5	892	5 ABB98134	ABB98134 Human PMM
33	1494.5	50.5	892	8 ADS10951	ADS10951 Human the
34	1494.5	50.5	969	5 AAE24168	AAE24168 Human dip
35	1490.5	50.4	581	5 ADR41521	ADR41521 Human CD-
36	1488.5	50.3	869	5 AAE24169	AAE24169 Alternati
37	1486.5	50.3	863	5 ADI16690	ADI16690 Human NOV
38	1486.5	50.3	863	5 ADI16688	ADI16688 Human NOV
39	1486.5	50.3	863	8 ADM42344	ADM42344 Human nov
40	1474.5	49.9	863	8 ADM42342	ADM42342 Human nov
41	1445	48.9	879	5 ABG61607	ABG61607 Human DPR
42	1445	48.9	879	5 ABG61608	ABG61608 Human DPR
43	1445	48.9	879	8 ADS10953	ADS10953 Human the
44	1412	47.8	597	7 ADE08485	ADE08485 Novel pro
45	1377.5	46.6	310	4 AAB47188	AAB47188 Human DPP

ALIGNMENTS

RESULT 1  
AAB47189  
ID AAB47189 standard; protein; 465 AA.  
XX  
AC AAB47189;  
XX  
DT 29-JUN-2001 (first entry)  
XX  
DE Human DPP8 318Thr-658Val+759Ala-882Ile.  
XX  
KW Human; dipeptidyl aminopeptidase; DPP8; prolol oligopeptidase;  
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;  
KW growth hormone deficiency; glucose level; mucosal regeneration;  
KW non-insulin dependent diabetes mellitus; glucose intolerance;  
KW immunosuppression.  
XX  
OS Homo sapiens.  
XX  
PN WO200119866-A1.  
XX  
PD 22-MAR-2001.  
XX  
PF 11-SEP-2000; 2000WO-AU001085.  
XX  
PR 10-SEP-1999; 99AU-00002762.  
PR 18-FEB-2000; 2000AU-00005709.  
XX  
PA (UNSY ) UNIV SYDNEY.  
XX  
PI Abbott CA, Gorell MD;  
XX WPI; 2001-281520/29.  
DR N-PSDB; AAC85696.

XX New human dipeptidyl aminopeptidase (DPP8) useful for cleaving  
PT substrates, identifying inhibitors of DPP8 catalytic activity which have  
FI therapeutic uses, and for detecting activated T cells.

PS Claim 5; Page 71-72; 78pp; English.

XX The sequences given in ABA47188-90 represent fragments of human  
CC dipeptidyl aminopeptidase (DPP8). DPP8 has substrate specificity for H-  
CC Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a prolyl  
CC oligopeptidase and a dipeptidyl peptidase, because it is capable of  
CC hydrolysing the peptide bond C-terminal to Pro in each of these  
CC compounds. DPP8 is homologous with human DPP4. DPP8 is useful for  
CC cleaving a substrate, and for detecting an activated T cell which  
CC involves measuring the level of DPP8 gene expression in a T cell. The  
CC level of DPP8 expression is detected by detecting the amount of DPP8 RNA  
CC in the cell. It is also useful for identifying a molecule capable of  
CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as  
CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,  
CC growth hormone deficiency, lowering glucose levels in non-insulin  
CC dependent diabetes mellitus and other disorders involving glucose  
CC intolerance, enhancing mucosal regeneration and as immunosuppressants

XX Sequence 465 AA;

Alignment Scores:  
Pred. No.: 1,21e-254 Length: 465  
Score: 2482.00 Matches: 465  
Percent Similarity: 100.0% Conservatives: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 83.9% Indels: 0  
DB: 4 Gaps: 0

US-10-825-632-6 (1-1669) x ABA47189 (1-465)

Qy 2 ACAGGTACAGCAAACTCTAAAGTCTCTTTAAGATGTCAGAAATAATGATGCTGAA 61  
Db 1 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 20  
Qy 62 GGAAGGATCATAGATGTCATAGATAAGAACTAATTCACACCTTTTGAGATTCTATTGAA 121  
Db 21 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 40  
Qy 122 GGAGTTGAATATATGCCAGACTGGATGGACTCTGAGGGAATAATGCTGGTCCATC 181  
Db 41 GlyValGluLysIleAlaArgAlaGlyTrpThrProGluGlyLysIleAlaTrpSerIle 60  
Qy 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATC 241  
Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 80  
Qy 242 CCAGTAGAAGATGATGTTATGTAAGAGCAGAGACTCATTGAGTCAGTGCCTGATCTGTG 301  
Db 81 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100  
Qy 302 ACGCCACTAATTATCTATGAGAACACACAGACATCTGGATATATATCCATGACATCTTT 361  
Db 101 ThrProLeuIleIleLysGluGlnThrAspIleTrpIleAsnIleHisAspIlePhe 120  
Qy 362 CATGTTTTCCCAAGTCACAGAGAAATTTAGATTATTTTGGCTCTGATGCAAA 421  
Db 121 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 140  
Qy 422 ACAGGTTTCGGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACGA 481  
Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysIleArg 160  
Qy 482 TCCAGTGGTGGGCTCGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA 541  
Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 180  
Qy 542 ATTACCAAGTGGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAA 601

Db 181 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 200  
Qy 602 GTCAGAAAGCGTGATATATTTGAAGGCACCAAAAGACTCCCTTTAGAGCATCATCTGTAC 661  
Db 201 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 220  
Qy 662 GTAGTCAGTTACGTAATCTCGAGAGGTGACAAAGGCTGACTGACCGGTGGCTACTCAT 721  
Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240  
Qy 722 TCTTGTGTCATCAGTCAGACTGTGACTTCTTTATAAGTAAGTAAAGTAAACACCAAGAAT 781  
Db 241 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 260  
Qy 782 CCACACTGTGTGCTCTTTTACAAGCTATCAAGCTCTCAAGATGACCCCAACTTGCACAAACA 841  
Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluLysAspProThrCysLysThr 280  
Qy 842 AAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTCGACTATCTCTCCA 901  
Db 281 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300  
Qy 902 GAAATTTCTCTTTTGAAGTACTACTCGATTTACATTTGATGGATGCTCTCAAGCCT 961  
Db 301 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320  
Qy 962 CATGATCTACAGCTGGAAAGAAATATCTCTACTGTCTGTTCATATATATGTTGCTCTCAG 1021  
Db 321 HisAspLeuGlnProGlyLysIleTyrProThrValLeuPheIleTyrGlyGlyProGln 340  
Qy 1022 GTTGCTATTGCTGGGGCCCACTCATCTGTGGATCTTCTATGATACAGGATACACGGAA 1081  
Db 341 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 360  
Qy 1082 CGTTATATGGTCCACCTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATG 1141  
Db 361 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 380  
Qy 1142 CAAGCAGAAAGTTCCCTCTCGAACCAAACTGTTTACTGCTCTTACATGTTTCTCTGGAT 1201  
Db 381 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAsp 400  
Qy 1202 GAGAAATGTCATTTTGCACATACCATGATATTTACTAGTGTGTTTGTAGTGGGCTGGAAG 1261  
Db 401 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 420  
Qy 1262 CCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGAGTTCTTGAATCGGA 1321  
Db 421 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 440  
Qy 1322 GRACATTTAGAACTGCTCTTTTGCATCTTTTGCATCTCTCAAGAAACCTTGGATCAGCTATTGCT 1381  
Db 441 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 460  
Qy 1382 GCTCTAAAGTGATA 1396  
Db 461 AlaLeuLysValIle 465

RESULT 2  
ABB97362  
ID ABB97362 standard; protein; 724 AA.  
XX  
AC ABB97362;  
XX  
DT 27-JUN-2002 (first entry)  
XX  
DE Novel human protein SEQ ID NO: 630.  
XX  
KW Human; antianaemic; vulnery; antiinflammatory; immunomodulator;  
KW antiinfectivity; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
KW expressed sequence tag.



AC ABB97361;  
DT 27-JUN-2002. (first entry)  
XX  
DE Novel human protein SEQ ID NO: 629.  
XX  
KW Human; antinaemic; vulnery; antinflammatory; immunomodulator;  
KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
KW expressed sequence tag.  
XX  
OS Homo sapiens.  
XX  
FN WO200222660-A2.  
XX  
XX 21-MAR-2002.  
XX  
PF 10-SEP-2001; 2001WO-US026015.  
XX  
PR 11-SEP-2000; 2000US-00659671.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX  
XX WPI; 2002-292408/33.  
DR N-PSDB; ABN32547.  
XX

XX An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis.  
XX  
XX Example 2; SEQ ID NO 629; 509pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke  
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.  
CC rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a protein of the invention  
XX

XX  
SQ Sequence 782 AA;  
  
Alignment Scores:  
Pred. No.: 1,616-254 Length: 782  
Score: 2482.00 Matches: 465  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 83.9% Indels: 0  
DB: 5 Gaps: 0

US-10-825-632-6 (1-1669) x ABB97361 (1-782)  
QY 2 ACAGGTACAGCAATCCTAAAGTCACITTTAAGATGTCAGAAATATGATGCTGAA 61  
DB 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337  
QY 62 GGAAGATCATAGATGTCATAGATAGGAACCTAATCAACCTTTTGAGATTCATTGAA 121  
DB 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 357  
QY 122 GGATGTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGGAATATCTTGGTCCATC 181  
DB 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTrpAlaTrpSerIle 377  
QY 182 CTATAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATATTATTC 241  
DB 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397

QY 242 CCAGTAGAAGATCATGTTATGGAAGGCAGAGACTCATTGAGTCAGTCCCTGATTCTGTG 301  
DB 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACGCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361  
DB 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePhe 437  
QY 362 CATGTTTTTCCCAAGTACAGAGAGAAATTTAGTTTATTTTTCCTCTGAAATGCAAA 421  
DB 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGTTTCCTCATTTATACAAAATTACATCTATTTTAAAGGAAGCAATATAACGA 481  
DB 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
QY 482 TCCAGTGGTGGCTGCTCTCCAAAGTATTTCAAGTGTCTCTATCAAGAGAGATAGCA 541  
DB 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluLysIleAla 497  
QY 542 ATTACCAAGTGTGAATGGGAAGTCTTGGCCGGCATGGATCTATATATCAAGTTGATGAA 601  
DB 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTCAGAGGCTGTATATTTTGAAGCACCAAAAGACTCCCTTTAGAGCATCACTGTAC 661  
DB 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 537  
QY 662 GTAGTCAGTTAGCTAAATCCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCAT 721  
DB 538 ValValSerTyrValAsnProGlyValThrArgLeuThrAspArgGlyTyrSerHis 557  
QY 722 TCTTGTGTCATCAGTCATCTGTGACTTCTTTTATAAGTAAAGTATAGTAAACCAAGAAT 781  
DB 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
QY 782 CCACACTGTGTGCTCCCTTTACAAAGCTATCAAGTCTCTGAAGATGACCAACTTGCAGAAC 841  
DB 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
QY 842 AAGGAATTTTGGGCCACCACTTTGGATTGAGTTCAGCAGTCTCTCTCTGACTACTCTCCCA 901  
DB 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
QY 902 GAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGCCT 961  
DB 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
QY 962 CATGATCTACAGCTGGAAAGAAATATCCTACTCTGCTGCTCATATATGTTGCTCTCAG 1021  
DB 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyLysProGln 657  
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAA 1081  
DB 658 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 677  
QY 1082 CGTTATATGGGTACCCCTGACAGAAATGAAACAGGGCTATTACTTATGAGATCTGTGGCATG 1141  
DB 678 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 697  
QY 1142 CAAGCAGAAAGATTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCGGAT 1201  
DB 698 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 717  
QY 1202 GAGAATGTCCTATTTTGCATACACAGTATATTTACTGAGTTTTTTTAGTGGCTGGAAG 1261  
DB 718 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValAlaGlyLys 737  
QY 1262 CCATATGATTTACAGATCTATCTCCAGGAGACACAGCATTAAGAGTTCTCTGAAATCGGGA 1321  
DB 738 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 757  
QY 1322 GAACATTATGAACCTGATCTTTTGGACTACCTTCAAGAAAAACCTTGGATCACGTATTGCT 1381



Db 758 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 777  
QY 1382 GCTCTAAAGTGATA 1396  
Db 778 AlaLeuLysValIle 782

RESULT 4  
ABU92026  
ID ABU92026 standard; protein; 831 AA.  
XX  
AC ABU92026;  
XX  
DT 15-JUL-2003 (first entry)  
XX  
DE Human protein modification and maintenance molecule-6 (PMWM-6).  
XX  
KW Human; protein modification and maintenance molecule; PMWM; cancer;  
KW cell proliferation disorder; atherosclerosis; neurological disorder;  
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;  
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;  
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;  
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;  
KW antiinflammatory; thyromimetic.  
XX  
OS Homo sapiens.  
XX  
PN WO2003031939-A2.  
XX  
PD 17-APR-2003.  
XX  
PF 11-OCT-2002; 2002WO-US032850.  
XX  
PR 12-OCT-2001; 2001US-0329689P.  
PR 25-OCT-2001; 2001US-0335703P.  
PR 09-NOV-2001; 2001US-0348887P.  
PR 28-NOV-2001; 2001US-0334145P.  
PR 06-DEC-2001; 2001US-0337451P.  
PR 14-DEC-2001; 2001US-0340584P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;  
PI Sprague WW, Hatalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;  
PI Marquis JP, Yao MG, Richardson TW, Tang TX, Jin P, Chien D;  
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
XX  
DR WPI; 2003-430274/40.  
DR N-PSDB; ACA92421.  
XX  
PT New human protein modification and maintenance molecules (PMWM), useful  
PT for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant PMWM expression e.g. cancer, atherosclerosis, or  
PT infections.  
XX  
PS Claim 1; Page 242-243; 311pp; English.  
XX

The present invention relates to the isolation of human protein  
modification and maintenance molecules (PMWM), and the polynucleotide  
sequences encoding them. A total of 40 PMWM polypeptides (designated PMWM  
-1 to PMWM-40) are disclosed. The sequences of the invention are useful  
for diagnosing a condition or disease associated with the expression of  
PMWM in a subject, preparing a polyclonal or monoclonal antibody, and  
generating an expression profile of a sample containing the  
polynucleotides. The diseases or conditions associated with decreased  
expression or overexpression of PMWM are cell proliferation disorders  
(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,  
allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
syndrome), gastrointestinal or epithelial disorders, and infections. The  
PMWM polypeptides or their fragments are useful in screening compounds

CC for effectiveness as agonists or antagonists of the polypeptides, or in  
CC altering the expression of the target polynucleotide and compounds that  
CC specifically bind to, or modulate the activity of the polypeptide.  
CC ABU92021-ABU92060 represent the human PMWM polypeptides of the invention  
XX  
SQ Sequence 831 AA;  
Alignment Scores:  
Pred. No.: 7,91e-251 Length: 831  
Score: 2447.50 Matches: 465  
Percent Similarity: 90.5% Conservative: 0  
Best Local Similarity: 90.5% Mismatches: 0  
Query Match: 82.8% Indels: 49  
DB: 6 Gaps: 1

US-10-825-632-6 (1-1669) x ABU92026 (1-831)

QY 2 ACAGGTACACCAATCTTAAGTCACCTTTTAAAGTGTCAAGTAATGATTGATCTGAA 61  
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337  
QY 62 GGAAGGATCATAGATGTATAGATAAGGAACCTAAATCAACCTTTTGAGATTCTATTGAA 121  
Db 338 GlyArgIleLeuAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 357  
QY 122 GGAGTTGAATATATTGCCAGAGCTGGATGCTCTGAGGAAATATGCTTGTGTCATC 181  
Db 358 GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIle 377  
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCCTACCTGAATTATTATC 241  
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
QY 242 CCAGTAGAAGATGATGTTATGAAAGGAGAGACTCATTGAGTCAGTCGCTGATTCTGTG 301  
Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACGCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361  
Db 418 ThrProLeuIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePhe 437  
QY 362 CATGTTTTTCCCAAGTCACGAAGAGAAATGAGTTTATTTTTCCTCTGAATGCATA 421  
Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGTTTCCGTCAATTTATACAAAATTAAGATCTATTATTAAGAAAGCAATATATAACGA 481  
Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
QY 482 TCCAGTGTGGGCTGCTGCTCCCAAGTGATTTTCAAGTGTCTTATCAAAAGAGAGATAGCA 541  
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluLeuAla 497  
QY 542 ATTTACAGTGTGAATGGGAAGTTCTTGCCGCGCATGGATCTAATATCAAGTTGATGAA 601  
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTCAGAAGCTGGTATATTTTGAAGCAGCAAGAGACTCCCTTTTAGACATCACCCTGTAC 661  
Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyr 537  
QY 662 GTAGTCAGTTACCTTAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACAT 721  
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgIleTyrSerHis 557  
QY 722 TCTTGTGTCATCAGTCAGCTGCTGCTCTTTTATAGTAAGTATAGTAAACAGAGAGAT 781  
Db 558 SerCysValSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
QY 782 CCACACTGTGTGCTCTTTTACAGCTATCAAGTCCTGAAGATGACCCCAACTTGCACAAACA 841  
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThr 597



Qy	362	CATGTTTTCCTCCCAAGTACGAGGAAATTCAGTTTATTTTGGCTCTGATCAAA	421	Qy	1295	CACAGCATAGAGTTCTCCTGAATCGGGAGAACATATGAACTGATCTTTTCACACTCTT	1354
Db	239	HieValPheProGlnSerHisGluGluGluGluPheAlaSerGluCysLys	258	Db	599	HisSerIleArgValProGlnSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeu	618
Qy	422	ACAGGTTTCCTGTCATTTATACAAATATACATCTATTTTAAAGGAAACCAATATAACGA	481	Qy	1355	CAAGAAACCTTCGATCAGCTATTGCTGCTCTTAAAGTGATA	1396
Db	259	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	278	Db	619	GlnGluAsnLeuGlySerArgIleAlaLeuLeuLysValIle	632
Qy	482	TCCAGTGGTGGCTGCTCCTCAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCA	541	RESULT 6			
Db	279	SerSerGlyGlyLeuProIleProSerAspPheLysCysProIleLysGluGluIleAla	298	ID	ADT04044	ADT04044 standard; protein; 738 AA.	
Qy	542	ATTACCAAGTGGTGAATGGGAAGTCTTGGCCGGCATGTCATATATCCAAAGTTGATGAA	601	AC	ADT04044;		
Db	299	IleThrSerGlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGlu	318	XX	30-DEC-2004 (first entry)		
Qy	602	GTCAAGAGCTGGTATATTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCCTGTAC	661	XX	Human protein modification and maintenance molecule protein SeqID24.		
Db	319	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr	338	XX	protein modification and maintenance molecule; PMMM; cytostatic;		
Qy	662	GTAGTCAGTTACCTAATCCTCGAGAGTGCACAGGCTGACTGACCGTGGCTACTCACAT	721	KW	immunomodulator; expression; immune disorder; cancer; human.		
Db	339	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgSerTyrSerHis	358	XX	Homo sapiens.		
Qy	722	TCTTGCTGCATCAGTCAGTCTGACTTCTTTATAGTATAGTAACTAGTAAACCAAGAAAT	781	PN	WO2004084806-A2.		
Db	359	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn	378	XX	07-OCT-2004.		
Qy	782	CCACACTGTGTGCTTTCACAGTATCAAGTCTCAAGTATGACCCCAACTGTCGAAACA	841	PD	16-MAR-2004; 2004WO-US008006.		
Db	379	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	398	PF	21-MAR-2003; 2003US-0456864P.		
Qy	842	AAGGATTTTGGCCACCATTTGGATTGATCAGAGTCTCTTCTGACTATATCTCTCCA	901	PR	03-APR-2003; 2003US-0460512P.		
Db	399	LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	418	PR	19-MAY-2003; 2003US-0472027P.		
Qy	902	GAATTTTCTCTTTTCAAGTACTACTGATTTACATTTGATGGGATGCTCTCAAGCCT	961	PR	30-MAY-2003; 2003US-0475072P.		
Db	419	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	438	PR	24-JUN-2003; 2003US-0482575P.		
Qy	962	CATGATCTACAGCTCGAAGAAATATCTACTGTGCTGTCATATATGTTGCTCTCAG	1021	PR	13-AUG-2003; 2003US-0495151P.		
Db	439	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	458	XX	(INCY-) INCYTE CORP.		
Qy	1021	-----	1021	PA	Baughn MR, Marquis JP, Kable AE, Chawla NK, Emerling BM, Lee SY;		
Db	459	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	478	PI	Hafalia AJA, Ramkumar J, Richardson TW, Wang JT, Nakamura LL;		
Qy	1021	-----	1021	PI	Yang J, Jin P, Becha SD, Wilson AD)		
Db	479	SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys	498	XX	WPI; 2004-710256/69.		
Qy	1022	-----GTTGCTATTGTCGGGCCCCAGTCACCTCTGTGG	1054	DR	N-PSDB; ADT04072.		
Db	499	PheGluGlyAlaPheLysTyrLysMetValAlaIleAlaGlyAlaProValThrLeuTrp	518	PT	New human protein modification and maintenance molecules (PMMM)		
Qy	1055	ATCTTCTATCATACAGATACACGACGTTATATGGGTCACTTCCCTGACCCAGATGACAG	1114	PT	polypeptide, useful in preparing a composition for treating a disease		
Db	519	IlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGln	538	PT	associated with decreased expression or overexpression of PMMM e.g.,		
Qy	1115	GCCTATTACTAGTCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGT	1174	XX	cancer.		
Db	539	GlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArg	558	PS	Claim 1; SEQ ID NO 24; 222pp; English/		
Qy	1175	TTACTGCTCTTACATGGTTTCTGGATGAGAAATGCCATTTTGACATACACGATATATTA	1234	CC	This invention relates to novel human protein modification and		
Db	559	LeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeu	578	CC	maintenance molecule (PMMM) proteins and the DNA sequences which encode		
Qy	1235	CTGAGTTTTCAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGAGA	1294	CC	them. The invention may be useful for the production of compounds with a		
Db	579	LeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArg	598	CC	cystostatic or immunomodulator activity. The proteins are useful in		
				CC	preparing a composition for diagnosing or treating a disease or condition		
				CC	associated with decreased expression or overexpression of functional		
				CC	PMMM, for example immune disorders or cancer. The present sequence is		
				CC	that of a human protein modification and maintenance molecule (PMMM)		
				CC	protein of the invention.		
				SQ	Sequence 738 AA;		
Alignment Scores:				Length: 738			
Pred. No.:				Matches: 458			
Score:				Conservative: 0			
Percent Similarity:				Mismatch: 6			
Best Local Similarity:				Indels: 1			
Query Match:				Gaps: 1			
DB:				8			

US-10-825-632-6 (1-1569) x ADT04044 (1-738)

QY 2 ACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATGTCAGAAATAATGATTGATGCTGAA 61  
DB 275 ThrSerThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 294

QY 62 GGAAGGATCATAGATCATAGATAGGAACCTAATCACTTTTGAAGATCTATTGAA 121  
DB 295 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 314

QY 122 GGGATTGAATATTCAGAGCTGGATGAGTCTCTGAGGAAAATATGCTTGCTCATC 191  
DB 315 GlyValGluTyrIleAlaArgAlaGlyIrrPmrProGluGlyLysTyrAlaIrrPserIle 334

QY 182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACCTGAAATATTATC 241  
DB 335 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 354

QY 242 CAGTAGAAGATGATGTTATGAAAGCGCAGAGACTCATTTAGTCAGTCCCTGATCTGTG 301  
DB 355 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 374

QY 302 ACSCCACTAATTATCATGAAGAAACAACAGACATCTGATAAAATATCCATGACATCTTT 361  
DB 375 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 394

QY 362 CATGTTTTCCCAAGTACCAAGAGGAAATTTAGATTTATTTTGGCTCTGAATCAAA 421  
DB 395 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 414

QY 422 ACAGGTTTCGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAATATAACGA 481  
DB 415 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysAsG 434

QY 482 TCCAGTGGTGGCTGCTCCAGTGTATTCAGTGTCTCTATCAAGAGGAGATAGCA 541  
DB 435 SerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 454

QY 542 ATTACAGTGGTGAATGGGAAGTCTTGGCGGCAATGATCTAATATCCAAAGTTGATGA 601  
DB 455 IleThrSerGlyLysIleValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 474

QY 602 GTACAGAGCTGGTATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTAC 661  
DB 475 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyr 494

QY 662 GTAGTCAGTTACGTAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAT 721  
DB 495 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 514

QY 722 TCTTGCTGCATCAGTCAGCTGTGACTCTTTTATAAGTAAAGTATAGTAAACCAAGAAAT 781  
DB 515 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 534

QY 782 CCACACTGTGTCTCCCTTTACAGCTATCAAGTCTCTGAAGTACCCCACTTGCAAAACA 841  
DB 535 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProLeuAla---Asn 553

QY 842 AAGGAATTTTGGCCCAACATTTGGATTTCAGCAGCTCTCTCTGCTACTATCTCTCCA 901  
DB 554 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 573

QY 902 GAATTTCTCTTTTGAAGTACTACTGGATTTACATGATGGATGCTCTACAGACT 961  
DB 574 GluIlePheSerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 593

QY 962 CATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTCTATATATATGCTGCTCAG 1021  
DB 594 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyLysProGln 613

QY 1022 GTTGCTATTGCTGGGGCCCAAGTCACTCTGFGGATCTTCTATGATACAGATACAGGAA 1081  
DB 614 ValAlaIleAlaGlyProProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 633

QY 1082 CGTTATATGCTCACCTGACCCAGCAATGAACAGGCTATTACTTAGGATCTGTGGCCATG 1141  
DB 634 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 653

QY 1142 CAAGCAGAAAGTTCCCTCTCTGAAACCAATCGCTTTACTGCTCTTACATGTTTCTCGAT 1201  
DB 654 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 673

QY 1202 GAGAAATGCTCATTTTGCACATACCACTAGTATATTACTAGTATTTTATGAGGCTCGAAAG 1261  
DB 674 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 693

QY 1262 CCATATGATTTCAGATCTCTCTCAGGAGAGACACAGATAAGAGTTTCTCGAATCGGGA 1321  
DB 694 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 713

QY 1322 GAACATTATGAACCTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTATTGCT 1381  
DB 714 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 733

QY 1382 GCTCTAAAAGTGATA 1396  
DB 734 AlaLeuLysValIle 738

RESULT 7  
ABU92030  
ID ABU92030 standard; protein; 824 AA..  
AC ABU92030;  
XX  
DT 15-JUL-2003 (first entry)  
DE Human protein modification and maintenance molecule-10 (PMMM-10).  
KW Human; protein modification and maintenance molecule; PMMM; cancer;  
KW cell proliferation disorder; atherosclerosis; neurological disorder;  
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;  
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;  
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;  
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;  
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnery;  
KW antiinflammatory; thyromimetic.  
OS Homo sapiens.  
XX  
XX WO2003031939-A2.  
XX  
PD 17-APR-2003.  
XX  
PF 11-OCT-2002; 2002WO-US032850..  
XX  
PR 12-OCT-2001; 2001US-0329689P.  
PR 25-OCT-2001; 2001US-0335703P.  
PR 09-NOV-2001; 2001US-0348887P.  
PR 28-NOV-2001; 2001US-0334145P.  
PR 06-DEC-2001; 2001US-0337451P.  
PR 14-DEC-2001; 2001US-0340584P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
PA  
XX Rankumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;  
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kabie AE, Yue H;  
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;  
PI Bhatia U, Burrill JD, Lee S, Blake JU, Ho A, Zheng W;  
XX WPI; 2003-430274/40.  
DR N-PSDB; ACA92425.  
XX  
XX New human protein modification and maintenance molecules (PMMM), useful  
PT for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or

```
PT infections.
XX Claim 1; Page 249-251; 31lpp; English.
XX
CC The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMM polypeptides (designated PMM
CC -1 to PMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PMM polypeptides of the invention
XX
SQ Sequence 824 AA;
      1
Alignment Scores:
Pred. No.:      4,11e-248      Length:      824
Score:          2422.00      Matches:      465
Percent Similarity: 82.3%      Conservative: 0
Best Local Similarity: 82.3%      Mismatches: 0
Query Match:      81.9%      Indels:      100
DB:              6              Gaps:      1

US-10-825-632-6 (1-1669) x ABU92030 (1-824)
QY      2 ACAGGTACAGCAATCTTAAGTCACTTTTAAGTGTGTCAGAAATAATGATTGCTGAA 61
      |||
DB      260 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 279
      |||
QY      62 GDAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCTATTGAA 121
      |||
DB      280 GlyArgIleLeuAspValIleAspLysGluLeuLeuGlnProPheGluLeuPheGlu 299
      |||
QY      122 GAGATTCAATATATCCAGACGTCGATGGACTCTCTGAGGAAATAATGCTTGTCATC 181
      |||
DB      300 GlyValGluIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 319
      |||
QY      182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATATTATC 241
      |||
DB      320 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 339
      |||
QY      242 CCAGTAGAAGATCATGTTATGAAAGGCGAGAGACTCATTGAGTCAGTGCCTGATTCTGTG 301
      |||
DB      340 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 359
      |||
QY      302 ACGCCAATATATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 361
      |||
DB      360 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 379
      |||
QY      362 CMTGTTTTCCCAAGTCACGAGGAAATGAGTTTATTTTGGCTCTCGATGCAA 421
      |||
DB      380 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 399
      |||
QY      422 AAGGTTTCCTCATTTATACAAATATACATCTATTTTAAAGGAAAGCAATATAACGA 481
      |||
DB      400 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 419
      |||
QY      482 TCCAGTGGTGGCTGCTCTCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA 541
      |||
DB      420 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 439
      |||
QY      542 ATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGATCTAATATCCAGTTGATGAA 601
      |||
```

```
DB      440 IleThrSerGlyGluTrpGluValValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 459
QY      602 GTCAGAGGCTGGTATATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCATCTGTAC 661
      |||
DB      460 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 479
      |||
QY      662 GTAGTCAGTTACCTAAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCAT 721
      |||
DB      480 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 499
      |||
QY      722 TCTTGCTGCATCAGTCAGCACTGTGACTTCTTTTAAAGTAAAGTATAGTAACCAAGAAT 781
      |||
DB      500 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 519
      |||
QY      782 CCACACTGTGTGCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAGAAC 841
      |||
DB      520 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 539
      |||
QY      842 AAGAAATTTGGCCACCATTTTGGATTCAGCAGAGTCTCTCTCTGACTATATCTCTCCA 901
      |||
DB      540 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 559
      |||
QY      902 GAAATTTCTCTTTTCAAGTACTACTGATTTTACATTTGATGGATCTCTCAAGCCT 961
      |||
DB      560 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 579
      |||
QY      962 CATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTTCATATATGTGGTCTCAG 1021
      |||
DB      580 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 599
      |||
QY      1021 ----- 1021
DB      600 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 619
      |||
QY      1021 ----- 1021
DB      620 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 639
      |||
QY      1021 ----- 1021
DB      640 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 659
      |||
QY      1021 ----- 1021
DB      660 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 679
      |||
QY      1021 ----- 1021
DB      680 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 699
      |||
QY      1022 GTTGCTATTGCTGGGCCCCAGTCACCTGTGGATCTTCTATGATACAGGATACACGGAA 1081
      |||
DB      700 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 719
      |||
QY      1082 CGTTATATGGTTCACCTGACAGATGAACAGGCTATTACTTAGGATCTGTGCCCATG 1141
      |||
DB      720 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 739
      |||
QY      1142 CAACAGCAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTTGAT 1201
      |||
DB      740 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAsp 759
      |||
QY      1202 GAGAAATCTCATTTTGCACATACCAGTATATTACTAGTATTTTGTAGTCAGGCTGGAAG 1261
      |||
DB      760 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 779
      |||
QY      1262 CCATATCATTTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTTCTTGAATCGGGA 1321
      |||
DB      780 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 799
      |||
QY      1322 GAAATATTAAGTGTGATCTTTTGTGACTACCTTCAAGAAACCTTGGATCAGTATTGCT 1381
      |||
DB      800 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 819
      |||
```

QY	1382	GCTCTAAAGTGATA	1396	Percent Similarity: 82.3%	Conservative: 0
DB	820	AlaLeuLysValIle	824	Best Local Similarity: 82.3%	Mismatches: 0
				Query Match: 81.9%	Indels: 100
				DB:	Gaps: 1
RESULT 8					
AAAB47187				US-10-825-632-6 (1-1669) x AAB47187 (1-882)	
ID	AAB47187	standard; protein; 882 AA.			
AC	AAB47187;				
XX					
DT	29-JUN-2001	(first entry)			
DE	Human DPP8.				
XX					
KW	Human, dipeptidyl aminopeptidase; DPP8; prollyl oligopeptidase;				
KW	dipeptidyl peptidase; DPP8; T cell; cleavage; diarrhoea;				
KW	growth hormone deficiency; glucose level; mucosal regeneration;				
KW	non-insulin dependent diabetes mellitus; glucose intolerance;				
KW	immunosuppression.				
OS	Homo sapiens.				
XX					
Key	Location/Qualifiers				
FT	Active-site	739			
FT	/note= "Forms part of Ser-Asp-His catalytic triad"	817			
FT	Active-site	849			
FT	/note= "Forms part of Ser-Asp-His catalytic triad"				
FT	Active-site	849			
FT	/note= "Forms part of Ser-Asp-His catalytic triad"				
XX					
XX	WO200119866-A1.				
XX					
XX	22-MAR-2001.				
XX					
XX	11-SEP-2000; 2000WO-AU001085.				
XX					
PR	10-SEP-1999; 99AU-00002762.				
PR	18-FEB-2000; 2000AU-00005709.				
XX					
XX	(UNSY ) UNIV SYDNEY.				
XX					
PI	Abbott CA, Gorell MD;				
XX					
DR	WPI; 2001-281520/29.				
DR	N-PSDB; AAC65694.				
XX					
PI	New human dipeptidyl aminopeptidase (DPP8) useful for cleaving				
FT	substrates, identifying inhibitors of DPP8 catalytic activity which have				
FT	therapeutic uses, and for detecting activated T cells.				
XX					
PS	Claim 1; Fig 2; 78pp; English.				
XX					
CC	This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has				
CC	substrate specificity for H-Ala-Pro-PNA, H-Gly-Pro-PNA and H-Arg-Pro-PNA.				
CC	Therefore, it is a prollyl oligopeptidase and a dipeptidyl peptidase,				
CC	because it is capable of hydrolysing the peptide bond C-terminal to Pro				
CC	in each of these compounds. DPP8 is homologous with human DPP8. DPP8 is				
CC	useful for cleaving a substrate, and for detecting an activated T cell				
CC	which involves measuring the level of DPP8 gene expression in a T cell.				
CC	The level of DPP8 expression is detected by detecting the amount of DPP8				
CC	RNA in the cell. It is also useful for identifying a molecule capable of				
CC	inhibiting DPP8 catalytic activity by DPP8. Molecules identified as				
CC	inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,				
CC	growth hormone deficiency, lowering glucose levels in non-insulin				
CC	dependent diabetes mellitus and other disorders involving glucose				
CC	intolerance, enhancing mucosal regeneration and as immunosuppressants				
XX					
SQ	Sequence 882 AA;				
Alignment Scores:					
Pred. No.:	4, 26e-248	Length:	882		
Score:	2422.00	Matches:	465		

```
Db 638 HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyProGln 657
Qy 1021 ----- 1021
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
Qy 1021 ----- 1021
Db 678 SerLeuGlyTyrValValValValLeuAspAsnArgGlySerCysHisArgGlyLeuLys 697
Qy 1021 ----- 1021
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluLeuAspGlnValGluGly 717
Qy 1021 ----- 1021
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
Qy 1021 ----- 1021
Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
Qy 1022 GTTGCTATTGCTGGGSCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAGGAA 1081
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
Qy 1082 CGTTATATGGTCCACCTCACAGCAATGAACAGGCTATTACTTAGGATCTGTGGCATG 1141
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
Qy 1142 CAAGCAGAAAGTCCCTCTCAACCAATCGTTTACTGCTCTTACATGGTTTCTCGAT 1201
Db 798 GlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
Qy 1202 GAGAATGTCATTGTCACATACAGTATATTACTGAGTTTTTTAGTGGGCTGGAAG 1261
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
Qy 1262 CMTATGATTACAGATCTATCTCAGGAGACACAGCATAAGAGTCTCTGATCGGGA 1321
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
Qy 1322 GRACATATCACTGATCTTTTGCATCTACCTCAAGAAACCTTGGATCAGCTATTGCT 1381
Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
Qy 1382 GCTCTAAAGTGATA 1396
Db 878 AlaLeuLysValIle 882
RESULT 9
ABG61591
ID ABG61591 standard; protein; 882 AA.
AC ABG61591;
XX ABG61591;
XX ABG61591;
DT 12-AUG-2002 (first entry)
DE Human DPPIV related serine protease DPPP-1.
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
OS Homo sapiens.
XX
XX WO200231134-A2.
XX
XX 18-APR-2002.
PD
```

```
XX 12-OCT-2001; 2001WO-US031874.
PF
XX
PR 12-OCT-2000; 2000US-0240117P.
XX
PA (FERR ) FERRING BV.
XX
PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX
DR WPI; 2002-444178/47.
DR N-PSDB; ABK33322.
XX
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
PS Claim 17; Fig 1; 113pp; English.
XX
CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPPP proteins
XX
SQ Sequence 882 AA;
Alignment Scores:
Pred. No.: 4,26e-248 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 5 Gaps: 1
US-10-825-632-6 (1-1669) x ABG61591 (1-882)
Qy 2 ACAGGTACAGCAATCTCTAAAGTCACCTTTAAGATGTCAGAAATTAATGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
Qy 62 GGAAGGATCATAGATGTCATAGATAGGAACATAATCAACCTTTTGAGATCTATTGAA 121
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIlePheGlu 357
Qy 122 GGAGTTGAATATATTCAGAGCTGGATGCTCTCTGAGGAGAAATATGCTTGTGTCATC 181
Db 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
Qy 182 CTACTAGATCGCTCCAGACTCGCTTACAGATAGTAGTGTGATCTCACCTGAAATTTATC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuLeuSerProGluLeuPheIle 397
Qy 242 CCAGTAGAAGATGATGTTATGGAAGGACAGACTCATTGAGTCCGCTGATTCGTG 301
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
Qy 302 ACGCCACTAATATCTATGAGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTT 361
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
Qy 362 CATGTTTTTCCCAAGTCCAGAGAGAAATGAGTTTATTTTTCCTCTGATGCAAA 421
Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIleAlaSerGluCysLys 457
```



QY 422 ACAGGTTTCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACGA 481  
Db 458 ThrGlyPheArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
QY 482 TCACAGTGGTGGGCTGCTCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCA 541  
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497  
QY 542 ATTACACAGTGGTCAATGGGAAGTCTTCGCGGCATGATCTAATATCCAAAGTGTGAA 601  
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 CTCAGAGAGCTGTATTTTGAAGCCACCAAGACTCCCTTTAGAGCATCCTGTAC 661  
Db 518 ValArgGluValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537  
QY 662 GTAGTCAGTACCTAATCCTCGAGAGGTGACAGGCTGACGCTGCTGCTACAT 721  
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
QY 722 TCTGTGTCATCAGTCAGCACTGTGACTTCTTTATAGTAAAGTAAAGTAAACCAAGAAT 781  
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
QY 782 CCACACTGTGTGCTTCCTTTACAGCTATCAAGTCTCAAGATGACCCAACTTCGCAAAACA 841  
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
QY 842 AAGGAATTTGGGCCACCACTTTGGATTGAGTTCAGCAGGTCTCTCTCTGACTATCTCTCCA 901  
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
QY 902 GAAATTTCTCTTTGAAAGTACTACTCGATTTCATTTGATGGATGCTCTCAAGCCT 961  
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
QY 962 CATGATCTACGCTGGAAGAAATCTCTACTGCTGTCTCATATATGCTGCTCTAG 1021  
Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657  
QY 1021 ----- 1021  
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677  
QY 1021 ----- 1021  
Db 678 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
QY 1021 ----- 1021  
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
QY 1021 ----- 1021  
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021  
Db 738 TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTGCTATTGCTGGGCCCAAGTCACTCTGGGATCTTCTATCATACAGATACAGGAA 1081  
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777  
QY 1082 CGTTATATGGTCCACCTGACCAAGTGAACAGGGCTATTACTAGGATCTGGGCCATG 1141  
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrLeuGlySerValAlaMet 797  
QY 1142 CAAGCAGAAAGTTCCTCTGAAACCAATCGTTACTGCTCTATCTGTTCTCGAT 1201  
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817  
QY 1202 GAGAATGTCCTATTTGGACATACCAAGTATATTACTGAGTGTCTTTAGTGAGGCGTGAAG 1261

Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
QY 1262 CCATATGATTTACAGATCTCTCTCAGAGAGACACACAGCATAGAGTTCCTGAATCGGA 1321  
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857  
QY 1322 GAACATTATGAACATGATCTTTTGCCTACTCTTCAAGAAACCTTGGATCAGTATTGCT 1381  
Db 858 GluHisTyrGlnLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877  
QY 1382 GCTCTAAAGTGATA 1396  
Db 878 AlaLeuLysValIle 882  
RESULT 10  
ID AAG78415 standard; protein; 882 AA.  
XX  
AC AAG78415;  
XX  
DT 12-APR-2002 (first entry)  
XX  
DE Amino acid sequence of 21953 human prollyl oligopeptidase.  
XX  
KW 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;  
KW cardiovascular disease; autoimmune disease; atopic allergy;  
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;  
KW antidiabetic; antiarthritic; antiseizmatic; antiinflammatory;  
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;  
KW Grave's disease; neuronal disorder; demyelinating disease.  
XX  
OS Homo sapiens.  
XX  
FN WO200179473-A2.  
XX  
PD 25-OCT-2001.  
XX  
EF 11-APR-2001; 2001WO-US040483.  
XX  
PR 18-APR-2001; 2000US-0197508P.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Meyers RA, Williamson M;  
XX  
DR WPI; 2002-034353/04.  
XX  
DR N-PSDB; AAH99934.  
PT New polypeptides 21953, member of human prollyl oligopeptidase family,  
PT useful as diagnostic targets and therapeutic agents for controlling  
PT cancer, lymphoma and leukemia.  
XX  
PS Claim 1; Page 102-103; 121pp; English.  
XX  
CC This invention relates to an isolated 21953 human prollyl oligopeptidase.  
CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,  
CC antichryoid, dermatological, antipsoriatic, antiaesthetic,  
CC ophthalmological, antiinflammatory, nontropic, antiparkinsonian,  
CC anticonvulsant, gynaecological, vasotropic, antianginal, cardiac,  
CC anticholesterolic, anorectic and metabolic in its action. Uses include  
CC gene therapy, expression or activity of 21953 protein modulator, it is  
CC useful for identifying a compound which binds to it and can be used in  
CC differentiating or detecting a cellular proliferative or  
CC differentiative disorder. The 21953 molecules can act as novel diagnostic  
CC targets and therapeutic agents for controlling disorders associated with  
CC the aberrant activity or degradation of peptide hormones e.g., disorders  
CC associated with cell differentiation and proliferation such as cancer,  
CC immune function, reproductive, neurological and cardiovascular cancer,  
CC The 21953 molecules are thus useful for treating and preventing cellular  
CC proliferative and differentiative disorders, haematopoietic neoplastic  
CC disorders, immune disorders such as autoimmune diseases, diabetes  
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,

CC	neural disorders, demyelinating diseases, vascular disorders and
CC	metabolism or pain disorders. This sequence represents the amino acid
CC	sequence of 21953 human prollyl oligopeptidase
XX	
SQ	Sequence 882 AA;
Alignment Scores:	
Pred. No.:	4,26e-248
Score:	2422.00
Percent Similarity:	82.3%
Best Local Similarity:	82.3%
Query Match:	81.9%
DB:	5
US-10-825-632-6 (1-1669) x AAG78415 (1-882)	
QY	2 AGAGGTACAGCAATCTTAAGTACCTTTAAGATGTCAGAAATATGATGCTGAA 61
DB	318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
QY	62 GGAAGATCATAGATGTCATAGATAAGAACTAATCAACCTTTGAGATTTCTATTGAA 121
DB	338 GlyArgIleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGlu 357
QY	122 GGAAGTGAATATATTCCAGAGCTGGATGGACTCCTCGAGGGAATAATATGCTGGTCCATC 181
DB	358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY	182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACTGATTAATTATTC 241
DB	378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY	242 CAGTGAAGATGATGTTATGGAAGGAGAGACTCATTGACTCAGTCCCTGATTCCTG 301
DB	398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY	302 ACCCACTAATATCTATGAAGAAACACAGACTCTGGATAAATATCCATGACATCTTT 361
DB	418 ThrProLeuIleIleTyrGluIleThrThrAspIleIleIleAsnIleHisAspIlePhe 437
QY	362 CATGTTTTCCCAAGTCCAGAGCAAAATTCAGATTTATTTTGGCTCTGAATCAAA 421
DB	438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
QY	422 ACAGGTTCCGTCATTATACAAATATCATCTATTTTAAAGGAACAAATATAACGA 481
DB	458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY	482 TCAGTGGTGGGCTGCTCTCCAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA 541
DB	478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497
QY	542 ATTACAGTGGTGAATGGGAAGTTCTTGGCCGCAATGATCTAATATCCAAAGTTGATGAA 601
DB	498 IleThrSerGlyIleTyrGluIleValLeuGluArgHisGlySerAsnIleGlnValAspGlu 517
QY	602 GTGAGAGGCTGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTAC 661
DB	518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY	662 GTAGTCAGTTACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCAT 721
DB	538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY	722 TCTGTGTCATCAGTCAGACTGTGACTTCTTTATAGTAAGTATAGTAAACCAGAAAT 781
DB	558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY	782 CCACACTGTGTGCTTACAGCTATCAAGTCTGAGATGACCCCACTTGCAAAACA 841
DB	578 ProHisCysValSerLeuTyrLysLeuSerSerProGluIleAspAspProThrCysLysThr 597
QY	842 AAGGAATTTGGGCCCACTTTTGGATTTCAGAGGTCTCTCTCTGACTATATCTCTCCA 901

Db	598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
Qy	902 GAAATTTCTCTTTTAAAGTACTACTGGATTTACATTTGATGATGCTCTCAAGCCT 961
Db	618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
Qy	962 CATGATCTACAGCTCGAAAGAAATATCTACTCTGCTGTGTTTCATATATGTTGGTCTCAG 1021
Db	638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
Qy	1021 ----- 1021
Db	658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677
Qy	1021 ----- 1021
Db	678 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
Qy	1021 ----- 1021
Db	698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717
Qy	1021 ----- 1021
Db	718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737
Qy	1021 ----- 1021
Db	738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757
Qy	1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAA 1081
Db	758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777
Qy	1082 CGTTATATGGTCCACCTGACCAGAAATGACAGGGCTATTACTAGGATCTGTGGCCATG 1141
Db	778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797
Qy	1142 CAGCAGAAAGTTCCCTCTGACCAATCGTTTACTGCTCTTACATGGTTTCTCGAT 1201
Db	798 GluAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817
Qy	1202 GAGAATGCTCCATTTTGCACATACACAGTATATTACTAGTCTTTTGTAGTGGGCTGGAAG 1261
Db	818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837
Qy	1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACACATAGAGTTCTCGAATCGGGA 1321
Db	838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
Qy	1322 GAACATTATGAATGATCTTTTGCATCTTTTGCATCTTCAAGAAACCTTGGATCAGTATTGCT 1381
Db	858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877
Qy	1382 GCTCTAAAGTGATA 1396
Db	878 AlaLeuLysValIle 882
RESULT 11	
ID	AAE24170 standard; protein; 882 AA.
XX	AAE24170;
AC	AAE24170;
XX	23-SEP-2002 (first entry)
DT	Human dipeptidyl peptidase 8 (DPP8) protein.
DE	
XX	Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW	autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW	graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW	antiviral; enzyme.

XX OS Homo sapiens.  
XX PN WO200234900-A1.  
XX PD 02-MAY-2002.  
XX PF 29-OCT-2001; 2001WO-AU001388.  
XX PR 27-OCT-2000; 2000AU-00001078.  
XX PA (UNSY ) UNIV SYDNEY.  
XX PI Abbott CA, Gorrell MD;  
XX DR WPI; 2002-454646/48.  
XX DR N-PSDB; AAD38956.  
XX PT New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors  
XX of DPP catalytic activity, which may be employed to treat e.g. neoplasia,  
XX type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV  
XX infection.  
XX PS Example; Fig 1; 91pp; English.  
XX CC The present invention relates to dipeptidyl peptidase (DPP) proteins and  
XX polynucleotides encoding such proteins. The DPP peptides are useful for  
XX screening inhibitors of DPP catalytic activity. The inhibitors are useful  
XX for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft  
XX rejection and HIV (human immuno deficiency virus) infection. The present  
XX sequence is human DPP8 protein  
XX SQ Sequence 882 AA;  
  
Alignment Scores:  
Pred. No.: 4,26e-248 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: 5 Gaps: 1  
  
US-10-825-632-6 (1-1669) x AAE24170 (1-882)  
QY 2 ACAGTCAGCGMAATCCTTAAGTCATCTTTAAGATGTCAGAAATAATGATGCTGAA 61  
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetileAspAlaGlu 337  
QY 62 GGAAGGATCATAGATGTCATAGATAGGAACCTAATCAACCTTTTGAGATTCCTATTGAA 121  
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357  
QY 122 GGAGTTGAATATATTGCCAGAGCTGGATGGACTCCTGAGGAAATATGCTTGGTCCATC 181  
Db 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerile 377  
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTATTATC 241  
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheile 397  
QY 242 CCAGTAGAAGATGATGTTATGGAAGGACAGACTCATGAGTCAGTCCTGATTCCTG 301  
Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACGCCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCATGACATCTTT 361  
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleIleTrpIleAsnIleHieAspIlePhe 437  
QY 362 CATCTTTTCCGAAAGTCAGAGAGAAATGAGTTATTTTGGCTCTGCAATGCAAA 421  
Db 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGGTTCCGTCATTATTATCAAAATATACATCTATTTTAAAGGAAAGCAATATAACGA 481

Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
QY 482 TCCAGTGGTGGCTGCTCTCAAGTGATTTCAAGTGATCTCTATCAAGAGAGAGATAGCA 541  
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497  
QY 542 ATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTTAATATATCAAGTTGATGAA 601  
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTCAGAAGGCTGATATATTGAAGGCACCAAGACTCCCTTTTAGACATCACCCTGTAC 661  
Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537  
QY 662 GTAGTCAGTTACGTAATCTCGAGAGGTGACAAGCTGACTGACCGTGGCTACTCCAT 721  
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
QY 722 TCTTGCTGCATCAGTCAGCACGTGCTCTTATTAAGTAAGTATAGTAAACAGAGAAT 781  
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
QY 782 CCACACTGTGTCTCCCTTTACAAGCTATCAAGCTCTCGAAGATGACCCAACTTGCAAAACA 841  
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThr 597  
QY 842 AAGGAATTTGGCCACCACTTTTGATTCAGCAGGTCTCTTCTGACTATACCTCCCA 901  
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
QY 902 GAAATTTCTCTTTGAAAGTACTACTGGATTTTACATTGATGGGATGCTCTCAAGCCT 961  
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
QY 962 CATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTCATATATGTGGTCTCAG 1021  
Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657  
QY 1021 ----- 1021  
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677  
QY 1021 ----- 1021  
Db 678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
QY 1021 ----- 1021  
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
QY 1021 ----- 1021  
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021  
Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTGCTATTGCTGGGGCCCAAGTCATCTCTGTGGATCTTCTATGATACAGGATACAGGAA 1081  
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777  
QY 1082 CATTATATGGTCCACCTGACCAAGATGAAACAGGCTATTACTTAGGATCTCTGGCCATG 1141  
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797  
QY 1142 CAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATCGTTTCTGGAT 1201  
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817  
QY 1202 GAGATGTCCATTGTCACATACAGTATATTACTAGTTTTTTAGTAGGCTCGAAG 1261

Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
 QY 1262 CCAATATGATTTACAGATCTATCTCCAGGAGAGACACAGCATAAGAGTCTCTGAATCGGGA 1321  
 Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857  
 QY 1322 GAACATTATGAATCCTGATCTTTTGGACCTTACCTTCAAGAAAACCTTGGATCAGTATGCT 1381  
 Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877  
 QY 1382 GCTCTAAAGTCATA 1396  
 Db 878 AlaLeuLysValIle 882  
 RESULT 12  
 AAU74749  
 ID AAU74749 standard; protein; 882 AA.  
 XX  
 AC AAU74749;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE Human protease PRTS-9 protein sequence.  
 XX  
 KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
 KW cell proliferative disorder; developmental disorder; epilepsy;  
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
 KW reproductive disorder; endometriosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200198468-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PF 13-JUN-2001; 2001WO-US019178.  
 XX  
 PR 16-JUN-2000; 2000US-0212336P.  
 PR 22-JUN-2000; 2000US-0213955P.  
 PR 29-JUN-2000; 2000US-0215396P.  
 PR 07-JUL-2000; 2000US-0216821P.  
 PR 14-JUL-2000; 2000US-0218946P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;  
 PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
 PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YF, Walsh RT;  
 PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
 PI Kallick DA;  
 XX  
 DR WPI; 2002-090437/12.  
 DR N-PSDB; ABK12892.  
 XX  
 PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in  
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative  
 PT (e.g. cancer) disorders.  
 XX  
 PS Claim 1; Page 140-142; 177pp; English.  
 XX  
 CC The present invention relates to twenty one new human proteases, referred  
 CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the  
 CC invention are useful in the diagnosis, treatment and prevention of  
 CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's  
 CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial  
 CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency  
 CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.  
 CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,  
 CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's  
 CC disease and reproductive e.g. infertility and endometriosis disorders.

CC Numerous other examples of each disorder are given in the specification.  
 CC The present protein sequence represents the human protease PRTS-9 protein  
 CC of the invention  
 XX  
 SQ Sequence 882 AA;  
 Alignment Scores:  
 Pred. No.: 4,266-248 Length: 882  
 Score: 2422.00 Matches: 465  
 Percent Similarity: 82.3% Conservative: 0  
 Best Local Similarity: 82.3% Mismatches: 0  
 Query Match: 81.9% Indels: 100  
 DB: 5 Gaps: 1  
 US-10-825-632-6 (1-1669) x AAU74749 (1-882)  
 QY 2 ACAGGTACAGCAAACTCTAAAGTCACCTTTTAAGATGTGAGAAATAATGATTGCTGAA 61  
 Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337  
 QY 62 GGAAGGATCATAGATGTCATAGATAAGGAACCTAATTAACCTTTTGAGATTCATTGAA 121  
 Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357  
 QY 122 GGAATTGAATATATGCCAGAGCTGGATGGACTCTGAGGGGAAAATATGCTTGCTCATC 181  
 Db 358 GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaIrrSerIle 377  
 QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATATTTATC 241  
 Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
 QY 242 CCAGTAGAAGATGATTTATGAAAGGACAGAGACTCATGTAGTCTGAGTCTGCTGATCTGTG 301  
 Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
 QY 302 ACGCCACTAATATCTATGAAGAAACACAGACATCTCGTAAATATATCATGACATCTTT 361  
 Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePhe 437  
 QY 362 CATGTTTTTCCCAAGTCAAGAGGAAATTTAGTATTTATTTTTCCTCTGAAATGCAAA 421  
 Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457  
 QY 422 ACAGGTTCCGTATTTATACAAAATTAATCATCTATTTTAAAGGAAGCAATATAACGA 481  
 Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
 QY 482 TCCAGTGGTGGCTGCTCTCCAGTCAATTTCAAGTGTCTTATCAAGAGAGAGATAGCA 541  
 Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497  
 QY 542 ATTACACAGTGGTGAATGGGAAGTTCTTGCCGGCATGGATCTAATATATCAAGTTGATGAA 601  
 Db 498 IleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
 QY 602 GTCAGAGGCTGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGACATCAGCTGTAC 661  
 Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537  
 QY 662 GTAGTCAGTTACCTAATCTCGAGAGGGTGCACAGGCTGACCTGACCGTGGCTACTCACAT 721  
 Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
 QY 722 TCTTGCTGCATCAGTCAGCAGCTGTGACTCTTTTATAAGTAAAGTAAAGTAAAGTAAAGTAA 781  
 Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
 QY 782 CCACACTGTGTGCTTCTTACAGCTATCAAGCTATCAAGTCTCTGAGATGACCCCACTTGCAAAACA 841  
 Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
 QY 842 AAGAAATTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTCTGACTATATCTCTCCA 901

Db 598 LysGluPheTrpAlaThrLeuLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
 Qy 902 GAAATTTCTCTTTTGAAGTACTACTGGATTACATGATGGATGCTCTACAGCCT 961  
 Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuIleGlyMetLeuIleYrLysPro 637  
 Qy 962 CATGATCTACGCTGGAAGAAATATCTCTACTGCTGCTTATATATGCTGCTCTCAG 1021  
 Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657  
 Qy 1021 ----- 1021  
 Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677  
 Qy 1021 ----- 1021  
 Db 678 SerLeuGlyTyrValValValValLeuAspAsnArgGlySerCysHisArgGlyLeuLys 697  
 Qy 1021 ----- 1021  
 Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
 Qy 1021 ----- 1021  
 Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
 Qy 1021 ----- 1021  
 Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
 Qy 1022 GTTGCTATTGCTGGGCCCCAGTCACCTCTGCGATCTCTATCATACAGATACAGGAA 1081  
 Db 758 ValAlaIleAlaGlyAlaProValThrLeuIlePheIlePheIlePheIlePheIlePheIle 777  
 Qy 1082 CTTTATATGGTACCTCCACAGATGACAGATGACAGGGCTATTACTAGATCTGFGCCATG 1141  
 Db 778 ArgTyrMetGlyHisPheAspGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMet 797  
 Qy 1142 CAACAGAAAGTCCCTCTGAAACCAATGTTACTGCTTATACATGTTCTCTGAT 1201  
 Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAsp 817  
 Qy 1202 GAGAAATGCTCAATTTGACATACACAGTATATTTAGTGGGCTGGAAG 1261  
 Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
 Qy 1262 CCATATGATTTACAGATCTATCTCAGAGAGACACAGCATAAGATCTCTGAATCGGA 1321  
 Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857  
 Qy 1322 GAACATTATGAATGCTCTTTTGCATCTTTGCTTACCTCAAGAAACCTTGCATCAGTATTGCT 1381  
 Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877  
 Qy 1382 GCTCTAAAGTGATA 1396  
 Db 878 AlaLeuLysValIle 882  
 RESULT 13  
 ID AD117086  
 XX AD117086 standard; protein; 882 AA.  
 AC AD117086;  
 XX  
 DT 15-APR-2004 (first entry)  
 XX  
 DE Human NOVX protein homologue SeqID 622.  
 XX  
 KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;  
 KW inflammation; autoimmune disorder; allergy; blood disorder;  
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
 KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;

KW Alzheimer's disease; infection; str.  
 XX Homo sapiens.  
 OS WO200268649-A2.  
 PN  
 XX  
 PD 06-SEP-2002.  
 PF  
 XX  
 PR 31-JAN-2002; 2002WO-US002785.  
 PR 31-JAN-2001; 2001US-0265395P.  
 PR 31-JAN-2001; 2001US-0265412P.  
 PR 31-JAN-2001; 2001US-0265514P.  
 PR 31-JAN-2001; 2001US-0265517P.  
 PR 02-FEB-2001; 2001US-0266406P.  
 PR 05-FEB-2001; 2001US-0266767P.  
 PR 07-FEB-2001; 2001US-0266975P.  
 PR 07-FEB-2001; 2001US-0267057P.  
 PR 08-FEB-2001; 2001US-0267459P.  
 PR 09-FEB-2001; 2001US-0268232P.  
 PR 15-FEB-2001; 2001US-0268974P.  
 PR 26-FEB-2001; 2001US-0271664P.  
 PR 27-FEB-2001; 2001US-0271839P.  
 PR 27-FEB-2001; 2001US-0271853P.  
 PR 02-MAR-2001; 2001US-0272788P.  
 PR 02-MAR-2001; 2001US-0273046P.  
 PR 14-MAR-2001; 2001US-0275925P.  
 PR 14-MAR-2001; 2001US-0275947P.  
 PR 14-MAR-2001; 2001US-0275950P.  
 PR 14-MAR-2001; 2001US-0275989P.  
 PR 15-MAR-2001; 2001US-0276448P.  
 PR 15-MAR-2001; 2001US-0276450P.  
 PR 16-MAR-2001; 2001US-0276397P.  
 PR 16-MAR-2001; 2001US-0276768P.  
 PR 20-MAR-2001; 2001US-0278652P.  
 PR 26-MAR-2001; 2001US-0278757P.  
 PR 26-MAR-2001; 2001US-0278778P.  
 PR 29-MAR-2001; 2001US-0279882P.  
 PR 29-MAR-2001; 2001US-0279884P.  
 PR 30-MAR-2001; 2001US-0280147P.  
 PR 11-APR-2001; 2001US-0282992P.  
 PR 11-APR-2001; 2001US-0283083P.  
 PR 20-APR-2001; 2001US-0285133P.  
 PR 23-APR-2001; 2001US-0285749P.  
 PR 03-MAY-2001; 2001US-0288327P.  
 PR 03-MAY-2001; 2001US-0288504P.  
 PR 29-MAY-2001; 2001US-0294047P.  
 PR 30-MAY-2001; 2001US-0294473P.  
 PR 08-JUN-2001; 2001US-0296964P.  
 PR 18-JUN-2001; 2001US-0298959P.  
 PR 19-JUN-2001; 2001US-0299324P.  
 PR 13-AUG-2001; 2001US-0312020P.  
 PR 16-AUG-2001; 2001US-0312889P.  
 PR 16-AUG-2001; 2001US-0312908P.  
 PR 21-AUG-2001; 2001US-0313390P.  
 PR 28-AUG-2001; 2001US-0315470P.  
 PR 31-AUG-2001; 2001US-0316447P.  
 PR 07-SEP-2001; 2001US-0318115P.  
 PR 07-SEP-2001; 2001US-0318118P.  
 PR 12-SEP-2001; 2001US-0318740P.  
 PR 19-SEP-2001; 2001US-0323379P.  
 PR 18-OCT-2001; 2001US-0330245P.  
 PR 18-OCT-2001; 2001US-0330308P.  
 PR 14-NOV-2001; 2001US-0332701P.  
 (CURA-) CURAGEN CORP.  
 Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;  
 Li L, Gangolli EA, Padigaru M, Anderson DW, Rastelli L, Miller CE;  
 Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CBA;  
 Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CB;  
 WPI; 2002-706998/76.

XX New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.

XX Disclosure# SEQ ID NO 622; 1498pp; English.

XX This invention relates to a novel nucleic acids, and encoded polypeptides  
XX thereof, which have properties related to the stimulation of biochemical  
XX or physiological responses in a cell, tissue, organ or organism.  
XX Specifically, it refers to the use of biologically active fragments for  
XX diagnostic and prognostic assays and furthermore in the treatment of  
XX diverse pathological conditions. The present invention describes novel  
XX human and murine NOVX proteins, as well as methods to modulate their  
XX expression using antisense oligos, ribozymes and peptide nucleic acids.  
XX The NOVX polypeptides, polynucleotides and antibodies are useful in  
XX treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,  
XX atherosclerosis, cancer and diabetes. Furthermore, they may be used in  
XX treating or preventing diseases such as inflammation, autoimmune  
XX disorders, allergies, blood disorders, acquired immunodeficiency syndrome  
XX (AIDS), obesity, asthma, immunoglobulin (Ig) A nephropathy, cirrhosis,  
XX arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy  
XX and epilepsy. Accordingly, these molecules have many activities including  
XX cytotactic, cardiant, antiinflammatory, immunosuppressive, antiallergic,  
XX haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,  
XX antiasthmatic, nephrotropic, antiarthritic, hepatotropic,  
XX neuroprotective, nootropic, antibacterial, virucide, antiparasitic,  
XX relaxant and anticonvulsant. In addition, they are useful in screening  
XX assays to identify small molecules that modulate or inhibit, for example,  
XX neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
XX used as in chromosome mapping, tissue typing, preventive medicine and  
XX pharmacogenomics. This polypeptide is a homologue of a human NOVX protein  
XX of the invention.

XX Sequence 882 AA;

Alignment Scores:  
Pred. No.: 4,26e-248 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: 5 Gaps: 1

US-10-825-632-6 (1-1669) x ADI117086 (1-882)

QY 2 ACAGGTACAGCAAAATCCCTAAAGTCACTTTTAAGATGTCAGAAATATGATGCTGAA 61  
DB 318 ThrGlyThrAlaAsnProLysValThrPhelysMetSerGluIleMetIleAspAlaGlu 337  
QY 62 GDAAGATCATAGATGTCATAGATAAGCACTAATTCACCTTTGAGATTCCTATTGAA 121  
DB 338 GlyArgIleAspValIleAspLysGluLeuIleGlnProGluIleLeuPheGlu 357  
QY 122 GDAAGTGAATATATGCGCAGAGCTGGATGGACTCCTCGAGGAGAAATATGCTGGTCCATC 181  
DB 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377  
QY 182 CTAATAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCCTGATTTATTATC 241  
DB 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
QY 242 CAGTAGAAGATGATGTTATGAAAGGAGAGACTCATTGAGTCAGTGCCTGATTTCTGTG 301  
DB 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACGCCAATAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGATCTTT 361  
DB 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437  
QY 362 CAGTGTTCCTCCCAAGTCACGAGGAAATTTAGATTATTTTGGCTCTGATGCAAA 421

DB 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGGTTTCCGTCATTATATACAAAATTAACATCTATTATTAAGGAAAGCAAAATATAACGA 481  
DB 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
QY 482 TCAGTGTGGTGGCTGCTCCAGTGGATTCAGTGTCTCAAGTGTCTCAAGAGGAGATAGCA 541  
DB 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497  
QY 542 ATTACAGTGGTGAATGGAAAGTCTTGGCGCGCATCGATCTATATATCAAGTTGATGAA 601  
DB 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTCAGAGGCTGGTATATTTTCAAGGCACCAAGACTCCCTTTAGAGCATCACTGTGATC 661  
DB 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 537  
QY 662 GTAGTCAGTTAGTAAATCCTGGAGAGTGACAGGCTGACTGACCTGGCTACTCACAT 721  
DB 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
QY 722 TCTTGTGTCATCAGTCAGCTGTGACTCTTTTATAAGTAAAGTATAGTAAACCAAGAAAT 781  
DB 558 SerCysCysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsn 577  
QY 782 CCACACTGTGTCTCCTTTACAAGCTAATCAAGTCTCTGAAGTACCCCACTGGCAAAACA 841  
DB 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
QY 842 AAGGAATTTGGCCACCATTTTGGATTGACAGCTCTCTCTCTGACTATATCTCTCCCA 901  
DB 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
QY 902 GAAATTTCTCTTTGAAAGTACTACTGATTTTACATTTGATGGGATGCTCTCAAGCCT 961  
DB 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
QY 962 CATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTGTGCTCTCAG 1021  
DB 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657  
QY 1021 ----- 1021  
DB 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677  
QY 1021 ----- 1021  
DB 678 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
QY 1021 ----- 1021  
DB 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
QY 1021 ----- 1021  
DB 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021  
DB 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAA 1081  
DB 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777  
QY 1082 CGTTATATGGTCCACCTCACAGATGACAGGCTATTACTTAGGATCTGTGGCCATG 1141  
DB 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797  
QY 1142 CAAGCAGAAAGTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGAT 1201  
DB 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817

QY 1202 GAGATGTCATTTGACATACCAAGTATATTACTAGATTTTGTAGTGGCTGGAAG 1261  
Db 818 GUASNVAIHIEPHEALAHISThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
QY 1262 CCATATGATTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGAGTTCTCTGAATCGGA 1321  
Db 838 ProTyAspLeuGlnIleTyProGlnGluArgHisSerIleArgValProGluSerGly 857  
QY 1322 GAACATTATGAAGTGCATCTTTTGACATCTACCTTCAAGAAAACCTTGGATCACGTATTGCT 1381  
Db 858 GluHisTy-GluLeuHisLeuLeuHisTyLeuGlnGluLeuGlySerArgGlieAla 877  
QY 1382 GCTCTAAAAGTGATA 1396  
Db 878 AlaLeuLysValIle 882

RESULT 14  
ID ABU07720  
AC ABU07720 standard; protein; 882 AA.  
XX  
XX  
DT 19-MAY-2003 (first entry)  
XX  
DE Human serine protease HIPHUM46.  
XX  
KW Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;  
KW serine protease activity modulation; dipeptidyl peptidase activity;  
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;  
KW Alzheimer's disease; parasupranuclear palsy; Huntington's disease;  
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;  
KW irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;  
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;  
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;  
KW multiple sclerosis.  
XX  
OS Homo sapiens.

XX  
FH Key  
FT Region  
FT 259..260  
FT /note= "Paired glutamates of the beta propeller domain"  
FT Active-site  
FT 739  
FT /label= Catalytic\_serine\_residue  
FT Active-site  
FT 817  
FT /label= Catalytic\_aspartate\_residue  
FT Active-site  
FT 849  
FT /label= Catalytic\_histidine\_residue  
XX  
PN GB2374869-A.  
XX  
XX 30-OCT-2002.  
XX  
XX 22-JAN-2002; 2002GB-00001404.  
XX  
XX 23-JAN-2001; 2001GB-00001760.  
XX  
XX (GLAX ) GLAXO GROUP LTD.  
XX  
XX Edbrooke MR, Lewis AP;  
XX  
XX WPI; 2003-150703/15.  
XX  
XX N-PSDB; ABX12255.  
XX  
XX Identifying modulators of serine protease activity useful for treating  
XX musculoskeletal diseases, by contacting cell expressing a novel serine  
XX protease polypeptide with a compound and monitoring serine protease  
XX activity.  
XX  
XX Claim 10; Page 26-29; 38pp; English.  
XX  
XX The invention relates to a method of identifying a substance that  
XX modulates serine protease activity, comprising contacting a cell such as

CC a neuronal cell, lung cell, intestinal cell or a cell infected with a  
CC virus, expressing a serine protease polypeptide (HIPHUM 46), or its  
CC variant having dipeptidyl peptidase activity, or a serine protease  
CC isolated from the cell with a test substance, and monitoring for serine  
CC protease activity. The method is useful for identifying a substance that  
CC modulates serine protease activity. A modulator of the serine protease is  
CC useful in the manufacture of a medicament for treatment or prophylaxis of  
CC a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus  
CC infection, Alzheimer's disease, parasupranuclear palsy, myotonic  
CC dystrophy, Huntington's disease or amyotrophic lateral sclerosis.  
CC Additional disease that may be treated using modulators of the serine  
CC protease include malabsorption syndromes, irritable bowel syndrome, lung  
CC disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,  
CC rectal polyps, small bowel tumours, colorectal tumours, anaemia,  
CC dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple  
CC sclerosis. The present sequence represents the amino acid sequence of the  
CC human serine protease HIPHUM46  
XX

XX SQ Sequence 882 AA;

Alignment Scores:  
Pred. No.: 4,26e-248 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: Gaps: 1

US-10-825-632-6 (1-1669) x ABU07720 (1-882)

QY 2 ACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGCTGAA 61  
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337  
QY 62 GGAAGATCATAGATCATAGATAGCAACTAATTCACCTTTTGAGATCTATTGAA 121  
Db 338 GlyArgIleIleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGlu 357  
QY 122 GGAGTTGATATATTGCCAGAGCTGGATGAGTCTCTCAGGGAAAAATATGCTGGTCAATC 181  
Db 358 GlyValGluTyrlleAlaArgAlaGlyTrpThrProGluGlyLysTyrlleAlaTrpSerIle 377  
QY 182 CTACTAGATCGCTCCAGACTCGGCTTCAGATAGTGTGTGATCTCAGCTGAATTTATTC 241  
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
QY 242 CCAGTAGAAGATGATGTTTATGGAAGAGCAGAGACTCATTTGAGTCAGTCCCTGATTCTGTG 301  
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACGCCACTAATTTATCTATGAGAAACACAGACATCTGGAATAAATATCCATGACATCTTT 361  
Db 418 ThrProLeuIleIleTyrlleGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437  
QY 362 CATGTTTTTCCCAAGTCCAGAGAGAAATGAGTTTATTTTTCCTCTGATCAAAA 421  
Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheIleAspGlyLys 457  
QY 422 ACAGGTTTTCCCTCATTTATACAAATTTACATCTATTATAAGGAAAGCAAAATATAACGA 481  
Db 458 ThrGlyPheArgHisLeuTyrlleThrSerIleLeuLysGluSerLysTyrlleAspG 477  
QY 482 TCAGTGTGTGGCTGCTGCTCCAGAGTATTTCAAGTGTCTTCAAGTGTCTTCAAGAGAGATAGA 541  
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497  
QY 542 ATTACCATGCTGTAATCGGAGTCTTGGCCGCGATGGATCTAATATCCAAAGTTGATGAA 601  
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 CTCAGAGCTGCTGATATTTTGAAGGCCAACAAAGATCCCTCTTTAGAGCATCACCTGTAC 661  
Db 518 ValArgArgLeuValTyrllePheGluGlyThrLysAspSerProLeuGluHisLeuTyrlle 537



RESULT 15  
ADRA41398  
ID ADR41398 standard; protein; 587 AA.  
XX AC ADR41398;  
XX AC ADR41398;  
XX 07-OCT-2004 (first entry)  
XX DE Human CD-like molecule HDQBS11, SEQ ID NO:197.  
XX  
XX Human; CD-like molecule; cluster of differentiation; diagnosis;  
XX prevention; immune disorder; immunodeficiency; autoimmune disorder;  
XX blood-related disorder; haematological disorder; haemostatic disorder;  
XX thrombolytic disorder; hyperproliferative disorder; cancer; tumour;  
XX apoptotic disorder; cardiovascular disorder; respiratory disorder;  
XX angiogenic disorder; neovascularisation; neurological disorder;  
XX endocrine disorder; reproductive system disorder; infectious disease;  
XX gastrointestinal disorder; drug screening; tissue regeneration;  
XX chemotaxis; gene therapy; antibody therapy; drug targeting;  
XX chromosome mapping; forensic analysis; immunophenotyping; cytostatic;  
XX haemostatic; tranquiliser; vulnerary; antiinflammatory; nephrotropic;  
XX cardiant; antiallergic; anti-HIV; antirheumatic; antiarthritic;  
XX antipsoriatic; immunosuppressive; vasotropic; neuroprotective;  
XX antithyroid; thyromimetic; gynaecological; virucide; hepatotropic;  
XX antibacterial; dermatological; chromosome 15q22.  
XX  
XX Homo sapiens.  
OS  
XX WO200226930-A2.  
PN  
XX  
XX 04-APR-2002.  
PD  
XX  
XX 25-SEP-2001; 2001WO-US029838.  
PF  
XX  
XX 26-SEP-2000; 2000US-0235484P.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
FA  
XX  
XX Rosen CA, Birse CE;  
PI  
XX  
XX WPI: 2002-405050/43.  
DR  
XX  
XX N-PSDB; ADR41222.  
DR  
XX  
XX Novel polynucleotides and polypeptides useful for treating, preventing or  
PT ameliorating cardiovascular, renal, neurovascular, and autoimmune  
PT disorders.  
PT  
XX  
XX Claim 11; SEQ ID NO 197; 1243pp; English.  
PS  
XX  
XX The invention relates to 167 novel human CD (cluster of differentiation)-  
CC like molecules (ADR41388-ADR41563) and to cDNAs encoding them (seqid:11)-  
CC  
XX  
XX Sequence 587 AA;  
SQ  
Alignment Scores:  
Pred. No.: 5.07e-247 Length: 587  
Score: 2411.00 Matches: 463  
Percent Similarity: 82.2% Conservative: 0  
Best Local Similarity: 82.2% Mismatches: 0  
Query Match: 81.5% Indels: 100  
DB: 5 Gaps: 1  
US-10-825-632-6 (1-1669) x ADR41398 (1-587)  
Qy 8 ACAGCAATCTCAAGTCACTTTTAAAGTGTGAGAAATAATGATTGATGCTGAAGGAGG 67  
Db 25 ThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArg 44  
Qy 68 ATCATAGATCTCATAGTAAGGAACCTTAATTCACCTTTTGAGATTCTATTGAAGGAGTT 127  
Db 45 IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyVal 64  
Qy 128 GAATATATTCCAGAGCTGGATGCTCTCTGAGGGGAAATATGCTTGTCATCTACTACTA 187

662 GTAGTCAGTACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAT 721  
Qy  
Db ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
722 TCTTCTGTCATCAGTCAGCAGCTGTCAGTCTCTTTATAGTAAGTATAGTAACAGAGAAAT 781  
Qy  
Db SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
782 CCACACTGTGTCTCCCTTTACAAGCTATCAAGTCTCAAGTCAACCACTTGCAGAAACA 841  
Qy  
Db ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
842 AAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGTCCTCTCTCTGACTATATCTCTCCA 901  
Qy  
Db LysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
902 GAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTATGGGATGCTCTACAGCCT 961  
Qy  
Db GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
962 CATGATCTACAGCTCGGAAGAAATATCCTACTGCTGCTTCATATATGCTGCTCTCAG 1021  
Qy  
Db HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657  
1021 ----- 1021  
Qy  
Db ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677  
1021 ----- 1021  
Qy  
Db SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
1021 ----- 1021  
Qy  
Db PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
1021 ----- 1021  
Qy  
Db LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
1021 ----- 1021  
Qy  
Db TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
1022 GTTGTATGTCGGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACACGGAA 1081  
Qy  
Db ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777  
1082 CATTATATGGTCACTGACAGAAATCAACAGAGCTATTACTTAGGATCTGTGGCCATG 1141  
Qy  
Db ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797  
1142 CAAGCAAGAAAGTCCCTCTGAACCAATCGTTTACTGCTTTPACATGTTTCTCGAAT 1201  
Qy  
Db GlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAsp 817  
1202 GAGAAATGTCATTTGACATACAGTATATATCTAGTATTTTATGAGGCTGGAAG 1261  
Qy  
Db GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
1262 CATATGATTTACAGATCTATCTCAGAGAGACACAGCATAGAGTTCCTGAATCGGA 1321  
Qy  
Db ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857  
1322 GAACATTATGAATGATCTTTTGCATCTACCTTCAAGAAACCTTGGATCACCTATTGCT 1381  
Qy  
Db GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877  
1382 GCTCTAAAGTGATA 1396  
Qy  
Db AlaLeuLysValIle 882

Db 65 GluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTyrSerIleLeu 84  
QY 188 GATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATCCAGTA 247  
Db 85 AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal 104  
QY 248 GAAGATGATGTTATGAAAGGAGAGACTCATTCAGTCAGTCGCTGATCTCTGACGCCA 307  
Db 105 GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro 124  
QY 308 CTAATTATCTATGAAGAACACACAGACTCTGGATAAATATATCCATGACATCTTTCATGT 367  
Db 125 LeuIleIleTyrGluGlnThrAspIleTrpIleAsnIleHisAspIlePheHisVal 144  
QY 368 TTTCCCAAGTCACAGAGAAATTCAGTTTATTTTGCCTCTGAATCCAAACAGGT 427  
Db 145 PheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGly 164  
QY 428 TTCCTGTCATTTATACAAATATACATCTATTTTAAAGGAAAGCAAAATATAACGATCCAGT 487  
Db 165 PheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSer 184  
QY 488 GGTGGCTCCTGCTCCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTACC 547  
Db 185 GlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThr 204  
QY 548 AGTGTGTAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGA 607  
Db 205 SerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 224  
QY 608 AGCTGTGATATTTTAAAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTTGTC 667  
Db 225 ArgLeuValTyrPheGluGlyThrLysAspSerProLeuLysHisLeuTyrValVal 244  
QY 668 AGTTAGTAAATCCTGGAGAGGTGACAGGCTGACTCACCGTGGCTACTCACATCTTGC 727  
Db 245 SerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyLysSerHisSerCys 264  
QY 728 TGCATCAGTCAGCTGTCCTCTTATAGTAAGTATAGTAACAGAGATCCACAC 787  
Db 265 CysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHis 284  
QY 788 TGTGTGTCCTTTACAGCTATCAAGTCTGAGATGACCCCACTTGCAAAACAAAGGAA 847  
Db 285 CysValSerLeuTyrLysLeuSerSerProGluAspSerProThrCysLysThrLysGlu 304  
QY 848 TTTTGGCCACCATTTTGGATTTCAGAGGTCTCTTCCTGACTATATCTCTCCAGAAAT 907  
Db 305 PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIle 324  
QY 908 TTCTCTTTTAAAGTACTGATTTACATTTGATGGGATGCTCTACAGCCTCATGAT 967  
Db 325 PheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAsp 344  
QY 968 CTACAGCTGAAAGAAATATCTTACTGTCTGTCTATATATGTTGGTCTCTCAG----- 1021  
Db 345 LeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGln 364  
QY 1021 ----- 1021  
Db 365 LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeu 384  
QY 1021 ----- 1021  
Db 385 GlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu 404  
QY 1021 ----- 1021  
Db 405 GlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGln 424  
QY 1021 ----- 1021

Db 425 TyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSer 444  
QY 1022 -----GTTGCT 1027  
Db 445 TyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla 464  
QY 1028 ATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTTATGATACAGGATACACGGAACGTTAT 1087  
Db 465 IleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyr 484  
QY 1088 ATGGGTACCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCA 1147  
Db 485 MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla 504  
QY 1148 GAAAAGTTCCCTCTCAACCAAAATCGTTTACTGTCTTACATGGTTTCTCGATGAGAAAT 1207  
Db 505 GluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsn 524  
QY 1208 GTCCATTTTGCACATACCAAGTATATTACTGAGTTTTTTTAGTGAGGGCTGGAAAGCCATAT 1267  
Db 525 ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr 544  
QY 1268 GATTTACAGATCTATCTCAGGAGAGACACAGCATAAGAGTTCTCTGATCGGAGAACAT 1327  
Db 545 AspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis 564  
QY 1328 TATGAACTGCATCTTTTGCACTTACCTTCAAGAAACCTTGGATCAGGTATTGCTCTCTA 1387  
Db 565 TyrGluLeuHisLeuLeuHisTyrLeuGlnLeuGlySerArgIleAlaAlaLeu 584  
QY 1388 ARAAGTGATA 1396  
Db 585 LysValIle 587

Search completed: May 2, 2006, 01:02:50  
Job time : 290.372 secs

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 01:03:49 ; Search time 11.3267 Seconds  
(without alignments)  
2835.522 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 2957  
Sequence: 1 aacagctacgcaatctcta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+.n2p.model -DEV=xlp  
-O=/abs/ABSSWEB.spool/US10825632/runat\_01052006\_105948\_3262/app\_query.fasta\_1  
-DB=PIR -QFMT=FASTA -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p  
-USER=US10825632 @CGN 1 1 92 @runat\_01052006\_105948\_3262 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THRAADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

ALIGNMENTS

RESULT 1  
T32919  
hypothehtical protein K02F2.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T32919  
R;Maggi, L.; Goela, D.  
submitted to the EMBL Data Library, January 1998  
A;Description: The sequence of C. elegans cosmid K02F2.  
A;Reference number: Z21246  
A;Accession: T32919  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-931 <MAG>  
A;Cross-references: UNIPROT:O4987; UNIPARC:UPI0000076BD8; EMBL:AF043699; PIDN:AAB97564  
A;Experimental source: strain Bristol N2; clone K02F2  
C;Genetics:  
A;Gene: CBSP:K02F2.1  
A;Map position: 1  
A;Introns: 58/3; 82/2; 131/2; 178/2; 275/3; 322/2; 404/2; 441/1; 464/2; 486/3; 528/2; 5

Alignment Scores:		Pred. No.:	4.63e-31	Length:	931
Score:		Percent Similarity:	462.00	Matches:	154
Best Local Similarity:		Conservative:	38.7%	Conservative:	79
Query Match:		Best Local Similarity:	25.6%	Mismatches:	177
DB:		Query Match:	15.6%	Indels:	192
			2	Gaps:	23
US-10-825-632-6 (1-1669) x T32919 (1-931)					
Qy	5	GGTACAGCAAACTCTTAAGTACATCTTTTAAGATGTCAGAAAATAATGATTGCTGAAGGA	64		
Db	355	GlyThrLysAsnAlaTyrSerThrLeuArgMet-----ValIleLeuGluAsnGly	371		
Qy	65	AGGATCATGATGTCATAGATAAGAACTAATATTCACCTTTTGAGATCTTATTGTAAGGA	124		
Db	372	LysAlaTyrAspValProLeuLysAsp-----GluValIleTyrLysHis	386		

```
QY 125 -----GTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAAATATGCTTGG 175
Db 387 CysProPheTyrGluTyrIleThrArgAlaGlyPhePheSerAspGlyThrThrValTrp 406
QY 176 TCCATCCTACTAGATCGCTCCAGACCTCGCTACAGATAGTGTGATC----- 223
Db 407 ValGlnValMetSerArgAspGlnAlaGlnCysSerLeuLeuLeuIleProTyrThrAsp 426
QY 224 -----TCACCT 229
Db 427 PheLeuLeuProGluGluLeuGlyCysIleSerIleLeuAspAsnLeuLeuSerThr 446
QY 230 GAATATTATCCAGTAGAGATGATTATGGAAGCAGAGACTCATTCAGTCAGTG 289
Db 447 AspLeuAsnMetGlyValTrpAspAsp---LysSerHisGluGluThrMetGluLysPro 465
QY 290 CTGATTCCTGTGACG-----CCACTAATTATCTATGAAGAACAACAGACATCTGGATA 343
Db 466 ProArgGlyLysLeuArgGlyThrValGlnIleHisIleLysAlaArgAsnAspTyrTrpIle 485
QY 344 AATATCCATGATCTTTTCATGTTTTTCCCAAGAGTCACGAAGAGAA-----ATTGAG 397
Db 486 AsnThrHisAsnAlaIleTyrProLeuLysIleThrAspGluGluHisProMetTyrGlu 505
QY 398 TTTATTTTGGCTCTGTAATGCMAAACAGGTTTCCTGTCATTATACAAATATACATCTATT 457
Db 506 PheIleTyrCysLeuGluLysProAsnGly---SerCysLeuAlaLeuIleSerAlaGlu 524
QY 458 TTAAGAGAAAGCAATATAAACGATCCAGTGGTGGCTGCTGCTCCAGATGATTTCAAG 517
Db 525 LeuAspGlnAsnGlyTyr----- 530
QY 518 TGTCTCATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGAGTCTTGGCCGGCAT 577
Db 531 CysArgHisThrGluGluLysLeuLeuMetAlaGluAsnPheSerIle-----AsnLys 548
QY 578 GGATCTAATATCCAAAGTTGATGAAGTCAGAGCGCTGATATTTTGAAGGCACCAAGAC 637
Db 549 SerMetGlyIleValValAspGluValArgGluLeuValTyrTrpValAlaAsnGluSer 568
QY 638 TCCCTTTAGACATCACCTGTAGTACGTAGTCAGTTAGCTAGTAAATCCTGGAGAGGTGACAAG 697
Db 569 HisProThrGluTrpAsnIle---CysValSerHisTyrArgThrGlyGlnHisAlaGln 587
QY 698 CTGACTACCGTGGCTACTCACATTTCTGCTGCATCAGTCACACTGTGACTTCTTTATA 757
Db 588 LeuThrGluSerGlyIle-----Cys-----PheLys 596
QY 758 AGTAAGTATAGTAACCAAGAG-----AATCCACACTGTGTGTCCTCTTTACAAG 805
Db 597 SerGluArgAlaAsnGlyLysLeuAlaLeuAspLeuAspHisGlyPheAlaCysTyrMet 616
QY 806 CTATCA-----AGTCCTGAAGATGACCCCAACTTGCAAAACAAAGGAATTT---TGG--- 853
Db 617 ThrSerValGlySerProAlaGlu-----CysArgPheTyrSerPheArgTrpLys 633
QY 854 -----GCCACCAVTTTGGATTTCAGCAGGTCCT 880
Db 634 GluAsnGluValLeuProSerThrValTyrAlaAlaAsnIleThrValSerGlyHisPro 653
QY 881 ---CTTCCTGAC-----TATACTCTCCAGAAATTTTCTCTTTTGAAGT---ACTACT 928
Db 654 GlyGlnProAspLeuHisPheAspSerProGluMetIleGluPheGlnSerLysLeuThr 673
QY 929 GGATTTACATTTGATGGGATGCTTACAAAGCTCATGATCTACAGCCTCGAAGAAATAT 988
Db 674 GlyLeuMetHisTyrAlaMetIleLeuArgProSerAsnPheAspProTyrLysLysTyr 693
QY 989 CCTACTGTGCTGTATATATGTTGGTCTCT----- 1018
Db 694 ProValPheHisTyrValTyrGlyGlyProGlyIleGlnIleValHisAsnAspPheSer 713
```

```
QY 1018 ----- 1018
Db 714 TrpIleGlnTyrIleArgPheCysArgLeuGlyTyrValValValPheIleAspAsnArg 733
QY 1018 ----- 1018
Db 734 GlySerAlaHisArgGlyIleGluPheGluArgHisIleHisLysLysMetGlyThrVal 753
QY 1018 ----- 1018
Db 754 GluValGluAspGlnValGluGlyLeuGlnMetLeuAlaGluArgThrGlyGlyPheMet 773
QY 1018 ----- 1018
Db 774 AspMetSerArgValValValHisGlyTrpSerTyrGlyTyrMetAlaLeuGlnMet 793
QY 1019 -----CAGTTGCTATTGCTGGGGCCCCCACTCATCTCTGTGG 1054
Db 794 IleAlaLysHisProAsnIleTyrArgAlaAlaIleAlaGlyIleAlaValSerAspTrp 813
QY 1055 ATCTTCTATGATACAGGATACACGGAAGCTTATATGGTCCACCTGACCGAATGAACAG 1114
Db 814 ArgLeuTyrAspThrAlaTyrThrGluArgTyrMetGlyTyrPro---LeuGluGluHis 832
QY 1115 GGCTATTACTTAGGATCTGTGGCCATCAAGCAGAAAAGTTCCCTCTGACCAATCCT 1174
Db 833 ValIlyGlyAlaSerSerIleThrGlyLeuValGluLysLeuProAspGluProAsnArg 852
QY 1175 TTACTGCTCTTACATGGTTTCTTGATGAGATGTCCATTTTGCACATACCATATATTA 1234
Db 853 LeuMetLeuValHisGlyLeuMetAspGluAsnValHisPheAlaHisLeuThrHisLeu 872
QY 1235 CTGAGTTTTTTAGTGGGCTGGAAAGCCATATATGATTATACAGATCTATCTCAGGAGAGA 1294
Db 873 ValAspGluCysIleLysLysGlyLysTyrPheIleGluLeuValIlePheProAsnGluArg 892
QY 1295 CACAGCATAGAGTCTCTGAATCGGAGAAACATATTATGAATCGATCTTTTGCACACTCTT 1354
Db 893 HisGlyValArgAsnAsnAspAlaSerIleTyrLeuAspAlaArgMetTyrPheAla 912
QY 1355 CAAGAA 1360
Db 913 GlnGln 914

RESULT 2
A:87516
dipterydyl peptidase IV [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87516
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001.
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249, MUID:21173698; PMID:11259647
A:Accession: A87516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1738 <SFO>
A:Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI000000C7616; GB:AE005673; NID:g13423647; F
C:Genetic Code
A:Gene: CC2154

Alignment Scores:
Pred. No.: 2,59e-28 Length: 738
Score: 429.50 Matches: 123
Percent Similarity: 36.5% Conservative: 57
Best Local Similarity: 24.9% Mismatches: 146
Query Match: 14.5% Indels: 167
DB: Gaps: 12

US-10-825-632-6 (1-1669) x A87516 (1-738)
```

```
QY 131 TATATTCCACAGCTGGATGACTCTCTGAGGGAATATGCTTGCTCCATCTACTAGAT 190
Db 284 TyrValAlaArgValAlaTrpSerAlaAspGlyLysThrValTyValGlnArgLeuSer 303
QY 191 CGCTCCACAGCTCGCTCAGACATGAGTGTGATCTCACCTGAATATTATTATCCAGTAGAA 250
Db 304 ArgAspGlnLysThrLeuAspLeuLeuAlaPhe----- 314
QY 251 GATGATGTTATGGAAGGCAGACACTCATTTGAGTCAGTGCCTGATCTTGTCAGCCCACTA 310
Db 315 -----ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCATCATCTTTTCAT 364
QY 311 -----ATTTATCTATGAAGAAACAACAGACATCTGGATAAATATCATCATCTTTTCAT 364
Db 321 GlyLysThrIleLeuThrAspThrAspProHisPheIleGluValSerAsnAspPheArg 340
QY 365 GTTTTTCCCAAGTCACGAAGAGGAAATTTGATTTTATTTTTCCTCTGAAATGCAAAACA 424
Db 341 ProLeuThrAspGly-----ThrPheLeuTrpGlySerGlu---LysAsp 354
QY 425 GGTTCCTCGTCAATTTATACAAATATCATCTATTTTAAGGAAGCAATATAAAGCATCC 484
Db 355 GlyAsnGlnHisLeuTyArgTyAla-----Ala 364
QY 485 AGTGGTGGCTGCTGCTCCAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATT 544
Db 365 AspGlyLysLeuIleAla-----GlnIle 372
QY 545 ACCAGTGTGTAATCGGAAGTTCTTGGCCGCATGATCTAATATCAAGTTGATGAAGTC 604
Db 373 ThrLysGlyAspTrpProValIleGlyLeuGluGly-----ValAspGluAla 388
QY 605 AGAAGCTGTGTATTTTGAAGGCACCAAGATCCCTTTAGAGCATCACTGTACGTA 664
Db 389 ArgLysValAlaIlePheSerAlaSerIleAspThrProIleGluArgLeuTyGlu 408
QY 665 GTCAGTTACGTAAATCTCTGAGAGGTGACAAAGCTGACTGACCGTGGTCTACTACATCT 724
Db 409 ValSerTyAlaLysProGlyLysProLysAlaLeuThrSerAlaGlyGlyTrpTrpAla 428
QY 725 TGCTGCATCAGTCAGCACTGTGACTTCTTATAAGTAAGTATAGTAACCAAGAAATCCA 784
Db 429 AlaLysValAlaAspAsnGlyGlyAlaPheAlaGlyThrTySerAspProLysThrPro 448
QY 785 CACTGTGTGCTCTTACAGACTATCAAGTCCT-----GAAGTACACCA 829
Db 449 SerGlnThrAlaLeuTySerAlaAspGlyLysArgValArgTrpIleGluGluAsnLys 468
QY 830 ACTTGCACAAACAAGGAATTTGGCCACCATTGGAATTCAGCGCTCTCTCTCTGAC 889
Db 469 LeuAlaGluGlyHisProTyTrp-----ProTyAlaAlaAsnLeuProGln 484
QY 890 TATACTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTACATTGATGGGATG 949
Db 485 -----ProGluPheGlySerLeuLysAlaAspGlyGluThrLeuHisTyGlu 501
QY 950 CTCTACAGCTCTATGATCTACAGCTGGGAAAGAAATATCCTACTGCTGCTGCTATAT 1009
Db 502 IleLeuLysProIleGlyPheAspProAlaLysLysTyProAlaIleValSerValTy 521
QY 1010 GGTGTGCTCT----- 1018
Db 522 GlyGlyProHisAlaGlnArgValMetLysAsnTrpHisSerProSerGluArgThrTy 541
QY 1018 ----- 1018
Db 542 LeuGluAlaGlyTyValIlePheLysLeuAspAsnArgGlySerGlyAsnArgSerAla 561
QY 1018 ----- 1018
Db 562 LysPheMetAlaLeuAspArgLysLeuGlyThrValGluValGluAspGlnLeuLeu 581
```

```
QY 1018 ----- 1018
Db 582 GlyAlaLysPheLeuAlaSerGlnProTyValAspAlaAspLysLeuGlyValMetGly 601
QY 1019 -----CAG 1021
Db 602 TrpSerTyGlyGlyPheMetAlaLeuMetLeuLeuThrAlaGluAsnThrProPheLys 621
QY 1022 GTTCTCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAA 1081
Db 622 AlaGlyAlaAlaGlyAlaProProThrGluTrpSerLeuTyAspThrAlaTyThrGlu 641
QY 1082 CGTTATATGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGCCCATG 1141
Db 642 ArgTyMetGlyLysProAspGluAsnLysAlaGlyTyAlaTySerAspIleAsnAsn 661
QY 1142 CAAGCAGAAAGTCCCTCTGAACCAATCGTTTACTCTCTTACATGTTTCTCGAT 1201
Db 662 ArgIleAspLysLeu-----AlaProGlySerLeuLeuLeuLeuHisGlyMetAlaAsp 679
QY 1202 GAGATGTCCATTTTGCACATACCATGATATATTACTGAGTTTTTTAGTGAGGCTGGAAG 1261
Db 680 AsnAsnValIlePheGluAsnSerThrArgLeuMetAlaAlaLeuGlnArgLysAlaIle 699
QY 1262 CCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGC 1300
Db 700 LeuPheGluMetAlaMetTyProGlyGluArgHisSer 712
RESULT 3
JC5142
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Xanthomonas maltophilia
C:Species: Xanthomonas maltophilia
C>Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5142
R:Kobashima, T.; Ito, K.; Yoshimoto, T.
J. Biochem. 120, 1111-1117, 1996
A:Title: Dipeptidyl peptidase IV from Xanthomonas maltophilia: Sequencing and expression
A:Reference number: JC5142; MUID: 97164011; PMID: 9010758
A:Accession: JC5142
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <KAB>
A:Cross-references: UNIPROT:P95782; UNIPARC:UPI0000085F8A; DDBI:D83263; NID:gl753196; PI
C:Comment: This enzyme catalyzes the hydrolytic removal of amino-terminal dipeptidyl re:
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidyl-peptidase hydrolase; membrane bound
P:4-18/Domain: transmembrane #status predicted <TMM>
F:610/Active site: Ser #status predicted
F:685/Active site: Asp #status predicted
F:717/Active site: His #status predicted
Alignment Scores:
Pred. No.: 5, 93e-22 Length: 741
Score: 355.00 Matches: 123
Percent Similarity: 36.1% Conservative: 77
Best Local Similarity: 22.2% Mismatches: 164
Query Match: 12.0% Indels: 190
Gaps: 17
DB: 2
US-10-825-632-6 (1-1669) x JC5142 (1-741)
QY 5 GGTACAGCAAAATCTTAAAGTCACCTTTTAAGATGTCAGAAATATGATGATGCTGAAGA 64
Db 257 GlyAspAlaAsnValGlnValLysLeuGlyValIleSerProAlaGluGlnAlaGlnThr 276
QY 65 AGATCATAGATGTCATAGATAAGGAAGCACTAATCAACCTTTTGAGATCTCTATTGGAAGA 124
Db 277 GlnTrpIleAspLeuGlyLysGluGlnAspIle----- 287
QY 125 GTTGAATATATTCAGAGCTGGATGACTCTCTGAGGAGAAATATGCTTGGTCCATCCTTA 184
Db 288 -----TyIleAlaArgValAsnTrpArgAspProGlnHisLeuSerPheGln---Arg 304
```

```
QY 185 CTAGATCGCTCCAGACTCGCTACAGATAGTTGATCTCACCTGAATATTATTATCCCA 244
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 305 GlnSerArgAspGlnLysLeuAspLeuValGluValThrLeu----- 319
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 245 GTAGAAGATGATGTTATGGAAGAGCAGAGACTCATTGAGTCAGTCGCTGATCTGTGAGC 304
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 320 -----AlaSerAsnGlnGlnArg----- 325
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 305 CCACATAATTATCTATGAAGAACAACACAGACATCTGGATAAATATCCATGACATCTTTCAT 364
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 326 -----ValLeuAlaHisGlnThrSerProThrTyrPheLeuProLeuHisAsnSerLeuArg 343
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 365 GTTTTTCCCAAGATCAGAAAGAGAAATGAGTTATATTTTCCCTCTGCAATGCAAAACA 424
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 344 PheLeu-----AspAspGlySerIleLeuTyrPheSerGlu---ArgThr 357
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 425 GGTTCCTGCTATTTATACAAATATACATCTATTTTAAAGGAAGCAAAATATAAAGCATCC 484
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 358 GlyPheGlnHisLeuTyrArgGile-----AspSerLysGlyLysAlaAla 372
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 485 AGTGTGGGCTGCTGCTCCAAAGTATTCAAAGTGTTCCTATCAAAGAGGAGATAGCAATT 544
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 373 -----AlaLeu 374
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 545 ACCAGTGTGAATGGAAGTTCTGGCCGCATGGATCTAATATCCAAAGTTGATGAAGTC 604
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 375 ThrHisGlyAsnTyrPheVal-----AspGluLeuLeuAlaValAspGluLys 390
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 605 AGAAGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTA 664
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 391 AlaGlyLeuAlaTyrPheArgAlaGlyIleGluSerAlaArgGluSerGlnIleTyrAla 410
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 665 GTCAAGTTACGTAATCTCGAGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACTCT 724
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 411 ValPro---LeuGlnGlyGlyGlnProGlnArgLeuSerLysAlaProGlyMetHisSer 429
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 725 TGCTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACACAGAGAAATCCA 784
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 430 AlaSerPheAlaArgAsnAlaSerValTyrValAspSerTyrPheAsnAsnSerThrPro 449
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 785 CACTGTGTGTCCTTTACAAAGCTATCAAGTCTGGAAGATGACCCCACTTGCAAAACAAG 844
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 450 ProGlnIleGluLeuPheArgAlaAsnGly----- 459
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 845 GAATTTTGGGCCACCATTTTG-----GATTCAGCAGGTCTCTCTCTGACTATACr--- 895
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 460 GluLysIleAlaThrLeuValGluAsnAspLeuAlaAspProLysHisProTyrAlaArg 479
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 896 -----CCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACA 937
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 480 TyrArgGluAlaGlnArgProValGluPheGlyThrLeuThrAlaAlaAspGlyLysThr 499
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 938 ---TTGTATGGGATGCTTACAGCCTCATGATCTACAGCCTGGGAAGAAATATCTTACT 994
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 500 ProLeuAsnTyrSerValIleLysProAlaGlyPheAspProAlaLysArgTyrProVal 519
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 995 GTGCTGTTTCATATATGGTGGTCT----- 1018
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 520 AlaValTyrValTyrGlyGlyProAlaSerGlnThrValThrAspSerTyrProGlyArg 539
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 1018 ----- 1018
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 540 GlyAspHisLeuPheAsnGlnTyrLeuAlaGlnGlnGlyTyrValValPheSerLeuAsp 559
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 1018 ----- 1018
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 560 AsnArgGlyThrProArgArgGlyArgAspPheGlyAlaLeuTyrGlyLysGlnGly 579
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 1019 -----CAGTTGCT----- 1027
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 580 ThrValGluValAlaAspGlnLeuArgGlyValAlaTyrPheLeuLysGlnGlnProTyrVal 599
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 1027 ----- 1027
Db      |||      |||      |||      |||      |||      |||      |||      |||
```

```
Db 600 AspProAlaArgIleGlyValGlnGlyTyrPheSerAsnGlyGlyTyrMetThrLeuMetLeu 619
QY 1028 -----ATTCTGGGGCCCCAGTCACTCTCTGTGG 1054
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 620 LeuAlaLysAlaSerAspSerTyrAlaCysGlyValAlaGlyAlaProValThrAspTyrP 639
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 1055 ATCTTCTATGATACAGGATACACGGAACGTTATATGGTCAACCTGACCCAGAGATGAACAG 1114
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 640 GlyLeuTyrAspSerHisTyrThrGluArgTyrMetAspLeuProAlaArgAsnAspAla 659
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 1115 GCCTATTACTTAGGATCTGTGCCATCAACAGCAGAAAGTTCCCTCTCTGAACCAATCGT 1174
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 660 GlyTyrArgGluAlaArgValLeuThrHisIleGluGlyLeuArgSerPro----- 676
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 1175 TTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGCATACCATCATATTATTA 1234
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 677 LeuLeuLeuIleHisGlyMetAlaAspAspAsnValLeuPheThrAsnSerThrSerLeu 696
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 1235 CTGAGTGTTTTAGTGAGCGCTGGAAAGCCATATGATTACAGATCTATCTCAGGAGACA 1294
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 697 MetSerAlaLeuGlnLysArgGlyGlnProPheGluLeuMetThrTyrProGlyAlaLys 716
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 1295 CACACATAAGAGTTCTCTGAATCGGAGAAACATTATGAACCTG 1336
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 717 HisGlyLeuSerGlyAlaAspAlaLeuHisArgTyrArgVal 730
Db      |||      |||      |||      |||      |||      |||      |||      |||
RESULT 4
S66261
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
C:Species: Flavobacterium meningosepticum
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S66261
R:Kabaehima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
A:Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV
A:Reference numbers: S66261; MUID:95314307; PMID:7793970
A:Accession: S66261
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-711 <KAB>
A:Cross-references: UNIPROT:O47900; UNIPARC:UPI00000B2115; EMBL:D42121; NID:9577283; PID
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase
Alignment Scores:
Pred. No.: 2.63e-17 Length: 711
Score: 300.50 Matches: 101
Percent Similarity: 35.1% Conservative: 60
Best Local Similarity: 22.0% Mismatches: 141
Query Match: 10.2% Indels: 157
DB: 2 Gaps: 11
US-10-825-632-6 (1-1669) x S66261 (1-711)
QY 314 ATCTATGAAGAAACAACAGACATCTGGATAAAATATCCATGATCTTTTCATGTTTTCCTC 373
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 302 LeuPheThrGluThrAspAsnAlaTyrIleGluThrAspAsnLeuThrMetGluPhe--- 320
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 374 CAAAGTCACGAAGAGGAAATTGAGTTTATTTTGGCTCTGTAATGCCAAACAGGTTCCGT 433
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 321 -----LeuAspAspAsnSerPheLeuTyrPalaSerGlu---ArgAspGlyHisArg 336
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 434 CATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATATAACGATCCAGTGGTGG 493
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 337 HisLeuTyrTyrPyrAspAlaAlaGlyLysLeuLysLys----- 349
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 494 CTGCCTGCTCCAAGTGTATTTCAAGTGTCTCTATCAAGAGGAGATAGCATATACCAGTGGT 553
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 350 -----GlnValSerLysGly 354
Db      |||      |||      |||      |||      |||      |||      |||      |||
QY 554 GAATGGGAAGTCTCTCGCCGCATGATCTAATATCAAGTTGATGAATCAGAGGCTG 613
Db      |||      |||      |||      |||      |||      |||      |||      |||
```

355	AspTrpGluIleAsnTyrTyrglyTyAsnProLysThrLysGlu-----	370
614	GTATATTGTAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCAGT---	670
371	ValTyrIleGlnThrThrGluLysGlySerIleAsnLys-----ValValSerLys	387
671	---TACGTAAATCCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCATCTTGGC	727
388	LeuAsnIleAsnThrGlyLysThrGlnLeuLeuSerAsnAlaGluGlyAsnAsnSerAla	407
728	TGCATCAGTCAGCTGTGACTCTTTATAGTAAGTATAGTAACACAGAGAATCCACAC	787
408	AlaPheSerLysThrPheAsnTyrPheIleAsnThrSerSerThrAlaLysValProThr	427
788	TGTGTGTCCCTTTACAGCTATCAAGTCCTGAA-----GATGAC	826
428	LysTyrIleLeuLysAspAlaAsnGlyLysAspValLysGluLeuGlnAsnAspAsp	447
827	CCAACTTGCMAAACCAAGGAATTTTGGGCCACCATTGTAATTCAGCAGGCTCTCTCT	886
448	LeuLeuAsnLysLeuLysSer-----Asp	455
887	GACTATACTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATTTTACATTTGATGGG	946
456	AsnPheIleAlaLysGluPheIleThrIleProAsnAlaAlaGlyAspGlnMetAsnAla	475
947	ATGCTCTCAAGCCTCATGACTACACCTGGAAAGAAATATCTTACTGTGCTGTTCATA	1006
476	TrpMetIleLysProLysAsnPheAspProAlaLysLysTyProValPheMetPheGln	495
1007	TATGGTGGTCT-----CAGTTGCT-----	1027
496	TyrSerGlyProGlySerGlnGlnValAlaAsnSerTrpAspGlyGlyAsnGlyIleTrp	515
1027	-----	1027
516	PheAspMetLeuAlaGlnLysGlyTyrLeuValValCysValAspGlyArgGlyThrGly	535
1027	-----	1027
536	PheArgGlyThrLysTyrLysLysValThrTyrLysAsnLeuGlyLysTyrGluIleGlu	555
1027	-----	1027
556	AspGlnIleThrAlaAlaLysTrpLeuGlyAsnGlnSerTyrValAspLysSerArgile	575
1027	-----	1027
576	GlyIlePheGlyTrpSerTyrGlyGlyTyrMetAlaSerLeuAlaMetThrLysGlyAla	595
1028	-----ATTGCTGGGGCCCGACTCTCTGTGGATCTTCTTATGATACA	1069
596	AspValPheLysMetGlyIleAlaValAlaProValThrAsnTrpArgPheTyrAspSer	615
1070	GGATACACGAACGTTATATGGTCCACCTGACCAAGATCAACAGCGCTATTACTTAGA	1129
616	IleTyrThrGluArgPheLeuGlnThrProGlnGluAsnLysAspGlyTyrAspLeuAsn	635
1130	TCTGTGGCCATCAAGCAGAAAGTTCCCTCTGMAACCAATCGTTTACTGTCTTACAT	1189
636	SerProThrThrTyrAlaLysLeuLeu-----LysGlyLysPheLeuLeuIleHis	652
1190	GGTTTCCTGGATCAGAAATGCCATTTTGACATACACAGTATATTACTGAGTTTTTATG	1249
653	GlyThrAlaAspAsnValHisPheGlnAsnSerMetGluPheSerGluAlaLeuIle	672
1250	AGGGCTGGAAGCCATATGATTACAGATCTTCTTCAGGAGACACACAGCATATAAGATT	1309
673	GlnAsnLysLysGlnPheAspPheMetAlaTyrProAspLysAsnHisSerIleIleGly	692
1310	CCTGAATCGGGAGACATTATGAATCATCTTTTGCATCTTCTTGCATCTTCAAGAAACCTT	1366
693	GlyAsnThrArgProGlnLeuTyrGluLysMetThrAsnTyrIleLeuGluAsnLeu	711

## RESULT 5

dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (*Saccharomyces cerevisiae*)  
A30247  
N:Alternate names: protein YHR028C  
C:Species: *Saccharomyces cerevisiae*  
C>Date: 07-Jun-1990 #sequence\_revision 30-May-1997 #text\_change 09-Jul-2004  
C:Accession: S46780; A30107  
R:Du, Z.  
submitted to the EMBL Data Library, June 1994  
A:Description: The sequence of *S. cerevisiae* cosmid 8082.  
A:Reference number: S46773  
A:Accession: S46780  
A:Molecule type: DNA  
A:Residues: 1-818 <DUZ>  
A:Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031A5F; EMBL:UI0399; NID:g500689; PIR:P18962  
R:Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.  
J. Cell Biol. 108, 1363-1373, 1989  
A:Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an integral membrane protein,  
A:Reference number: A30107; MUID:89174971; PMID:2647766  
A:Accession: A30107  
A:Molecule type: DNA  
A:Residues: 1-82, 'H', 84-124, 'N', 126-181, 'LRLET', 189-199, 'N', 201-365, 'DFKGRKERF', 376-518  
A:Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484  
A>Note: the authors translated the codon ACC for residue 572 as Asn  
C:Genetics:  
A:Gene: SGD:DAP2; STE13; MIPS:YHR028C  
A:Cross-references: SGD:S0001070; MIPS:YHR028C  
A:Map position: 8R  
C:Superfamily: dipeptidyl-peptidase IV  
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vacuole-associated protein  
F:30-45/Domain: transmembrane #status predicted <TM>  
F:63,79,110,139,392,421/Binding site: carboxylate (Asn) (covalent) #status predicted

Alignment Scores:

Pred. No.	2,77e-15	Length:	818
Score:	277.00	Matches:	121
Percent Similarity:	32.2%	Conservative:	64
Best Local Similarity:	21.0%	Mismatches:	170
Query Match:	9.4%	Indels:	220
DB:	1	Gaps:	22

US-10-825-632-6 (1-1669) x A30107 (1-818)

Qy	143	GCTGGATGACCTCTGAGCGAAATATGCTTGGTCATCCTACTAGATCGCTCCAGACT	202
Db <td>251<td>AlaTrpTrpSerProThrGlyAspTyrlleualapheluleysileaspGlusluGluVal</td><td>270</td></td>	251 <td>AlaTrpTrpSerProThrGlyAspTyrlleualapheluleysileaspGlusluGluVal</td> <td>270</td>	AlaTrpTrpSerProThrGlyAspTyrlleualapheluleysileaspGlusluGluVal	270
Qy <td>203<td>CGCCTACAGATGTGTGCATCCACCTGAATTATTATCCCATAGAGAAGATGTTTG</td><td>262</td></td>	203 <td>CGCCTACAGATGTGTGCATCCACCTGAATTATTATCCCATAGAGAAGATGTTTG</td> <td>262</td>	CGCCTACAGATGTGTGCATCCACCTGAATTATTATCCCATAGAGAAGATGTTTG	262
Db <td>271<td>GlygluPheile-----ProfyTyrlValGlnAspGluIlysaspiletyr</td><td>286</td></td>	271 <td>GlygluPheile-----ProfyTyrlValGlnAspGluIlysaspiletyr</td> <td>286</td>	GlygluPheile-----ProfyTyrlValGlnAspGluIlysaspiletyr	286
Qy <td>263<td>GAAGCGACAGACTCAATTCAGTCAGTCGCTGATCTCTGAGCCACATAATTATCATGAA</td><td>322</td></td>	263 <td>GAAGCGACAGACTCAATTCAGTCAGTCGCTGATCTCTGAGCCACATAATTATCATGAA</td> <td>322</td>	GAAGCGACAGACTCAATTCAGTCAGTCGCTGATCTCTGAGCCACATAATTATCATGAA	322
Db <td>287<td>ProGluMetArgSeriIlys---TyrlProIysSerGlyThrPro-----Asn</td><td>301</td></td>	287 <td>ProGluMetArgSeriIlys---TyrlProIysSerGlyThrPro-----Asn</td> <td>301</td>	ProGluMetArgSeriIlys---TyrlProIysSerGlyThrPro-----Asn	301
Qy <td>323<td>GAACAACACAGATCTGGATA---AAATCCATCAC-----ATCTTCATGTTTTTCCC</td><td>373</td></td>	323 <td>GAACAACACAGATCTGGATA---AAATCCATCAC-----ATCTTCATGTTTTTCCC</td> <td>373</td>	GAACAACACAGATCTGGATA---AAATCCATCAC-----ATCTTCATGTTTTTCCC	373
Db <td>302<td>ProHisalagluLeuTrpValTyrlSerMetIysaspGlyThrSerPheHisProArglie</td><td>321</td></td>	302 <td>ProHisalagluLeuTrpValTyrlSerMetIysaspGlyThrSerPheHisProArglie</td> <td>321</td>	ProHisalagluLeuTrpValTyrlSerMetIysaspGlyThrSerPheHisProArglie	321
Qy <td>374<td>CMAAGTCACGACGAGAAATTGAGTTTATTTT-----</td><td>406</td></td>	374 <td>CMAAGTCACGACGAGAAATTGAGTTTATTTT-----</td> <td>406</td>	CMAAGTCACGACGAGAAATTGAGTTTATTTT-----	406
Db <td>322<td>SerGlyAsnIlysaspGlySerLeuleulleThrGluValThrTrpValGlyAsnGly</td><td>341</td></td>	322 <td>SerGlyAsnIlysaspGlySerLeuleulleThrGluValThrTrpValGlyAsnGly</td> <td>341</td>	SerGlyAsnIlysaspGlySerLeuleulleThrGluValThrTrpValGlyAsnGly	341
Qy <td>407<td>-----GCCCTCGAATGCAAAACAGGTTTCGTCATTTATACAAAATTACA</td><td>451</td></td>	407 <td>-----GCCCTCGAATGCAAAACAGGTTTCGTCATTTATACAAAATTACA</td> <td>451</td>	-----GCCCTCGAATGCAAAACAGGTTTCGTCATTTATACAAAATTACA	451
Db <td>342<td>AsnValIeuValIysThrThrAspArgSerSerAspileleuThrValPheIleuileasp</td><td>361</td></td>	342 <td>AsnValIeuValIysThrThrAspArgSerSerAspileleuThrValPheIleuileasp</td> <td>361</td>	AsnValIeuValIysThrThrAspArgSerSerAspileleuThrValPheIleuileasp	361
Qy <td>452<td>TCTATTTTAAGGAAGCAATATATAACGA-----TCCAGTGGTGGG-----</td><td>493</td></td>	452 <td>TCTATTTTAAGGAAGCAATATATAACGA-----TCCAGTGGTGGG-----</td> <td>493</td>	TCTATTTTAAGGAAGCAATATATAACGA-----TCCAGTGGTGGG-----	493
Db <td>362<td>ThrllealalyThrSerAsnValValarghsnglusSerSerAsnGlyIglylTrpTrpglu</td><td>381</td></td>	362 <td>ThrllealalyThrSerAsnValValarghsnglusSerSerAsnGlyIglylTrpTrpglu</td> <td>381</td>	ThrllealalyThrSerAsnValValarghsnglusSerSerAsnGlyIglylTrpTrpglu	381
Qy <td>494<td>-----CTGCCTCTCCAAGTGATTTCAAGTGCTCTATCAAAGAG</td><td>532</td></td>	494 <td>-----CTGCCTCTCCAAGTGATTTCAAGTGCTCTATCAAAGAG</td> <td>532</td>	-----CTGCCTCTCCAAGTGATTTCAAGTGCTCTATCAAAGAG	532





## Alignment Scores:

Pred. No.: 1,86e-14 Length: 931  
 Score: 267.50 Matches: 114  
 Percent Similarity: 33.0% Conservative: 63  
 Best Local Similarity: 21.3% Mismatches: 164  
 Query Match: 9.0% Indels: 195  
 DB: 2 Gaps: 18

US-10-825-632-6 (1-1669) x A49737 (1-931)

```

QY 221 ATCTCACTGAAATTTATTCACAGTAGAAGATGATTTATGGAAGCAGAGACTCATT 280
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 445 ILeSerProAspThrPhe-----ArgPheGluIleThrAspArgAsnSerLysIle 461
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 281 GAATCAGTGCCT-----GATTCGTGAGCCCACTAAATATCTATGAAGAACA 328
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 462 LeuAspValLysValTyAspIleProSerSerGlnMetLeuThrValArgAsnThrAsn 481
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 329 ACAGACATC-----TGATA-----AATATCCATGACATCTTTCATGTTTTCCCA 376
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 482 SerAsnLeuPheAsnGlyTrrIleGluLysThrLysAspIleLeuSerIleProProLys 501
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 377 AGTCACGAAGGAATTGAG-----TTTATTTTGCCTCTGATGCAAAACAGGTTTC 430
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 502 ProGluLeuLysArgMetAspTyArgIleAspIleHisAlaAspSerArgGlyPhe 521
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 431 CGTCATTTATACAAATATCATCTATTTTAAAGGAAGCAATATATACAGTCCAGTGGT 490
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 522 SerHisLeuPheTyTyProThrValPhe-----531
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 491 GGGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATACCAGT 550
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 532 -----AlaLysGluProIleGlnLeuThrLys 540
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 551 GGTGAATGGGAAGTTCTTGGCCGCAT-----GGATCTAATATCAAGTTGATGAAGTC 604
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 541 GlyAsnTrpGluValThrGlyAsnGlyIleValGlyTyArgGluThrAsp----- 558
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 605 AGAAGGTGGTATATTTGAAGGCACCAAGACTCCCTTTAGACATCACTGTACGTA 664
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 559 -----ThrIlePheThrAlaAsnGluIleGlyValMetSerGlnHisLeuTySer 576
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 665 CTCAGTTACGTA-----AATCTCGGAGAG 688
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 577 IleSerLeuThrAspSerThrThrGlnAsnThrPheGlnSerLeuGlnAsnPro----- 594
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 689 GTGACAGGCTGACTGACCTGGCTACTCACAATCTTGTGCTCATCAGTCAGCACTGTGAC 748
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 595 -----SerAspLysTyArgPhePheTyArgPheGluLeuSerSerSerAlaArg 610
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 749 TTCCTTTAAGTAGTAGTAGTAACCAAGAAATCCA-----784
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 611 TyrAlaIleSerLysLysLeuGlyProAspThrProIleLysValAlaGlyProLeuThr 630
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 785 -----CACTGTGCTCCCTTTACAAGCTATCAAGTCTGA 820
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 631 ArgValLeuAsnValAlaGluIleHisAspSerSerIleLeuGlnLeuThrLysAspGlu 650
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 821 GATGACCCCAACTTCGAAACCAAGAAATTTGGGCCCAACCATTTTGGATTCCAGGTCCT 880
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 651 -----LysPheLysGlu-----Lys 655
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 881 CTTCTGCTACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACT-----CGA 931
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 656 IleLysAsnTyAspLeuPro-----IleThrSerTyLysThrMetValLeuAspAspGly 674
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 932 TTTACATTTGATGGATGCTCTACAGCCCTCATGATCTACAGCTCGGAAAGAAATATCCT 991
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 675 ValGluIleAsnTyIleGluIlePheAlaAsnLeuAsnProLysLysTyPro 694
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 992 ACTGTGCTGTTTCATATATGGTGGTCT-----1018
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 695 IleLeuValAsnIleTyArgGlyProGlySerGlnThrPheThrThrLysSerSerLeu 714
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  
```

## RESULT 7

T41703  
 dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: T41703  
 R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 submitted to the EMBL Data Library, August 1998  
 A;Reference number: Z22011  
 A;Accession: T41703  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-793 <MUR>  
 A;Cross-references: UNIPROT:O14073; UNIPARC:UPI000013AB95; EMBL:AL031180; PIDN:CAA20138  
 A;Experimental source: strain 972h-; cosmid c2811 - chimeric  
 C;Genetics:  
 A;Gene: SPAC2E11.08  
 A;Map position: 1  
 C;Superfamily: dipeptidyl-peptidase IV

Alignment Scores:  
 Pred. No.: 1.96e-14 Length: 793  
 Score: 267.00 Matches: 100  
 Percent Similarity: 35.1% Conservative: 37  
 Best Local Similarity: 25.6% Mismatches: 127  
 Query Match: 9.0% Indels: 126  
 DB: 2 Gaps: 14



C;Comment: This protein is localized to the bile canaliculus, which is the apical domain  
C;Superfamily: dipeptidyl-peptidase IV  
F;1-792/Product: dipeptidyl-peptidase, membrane-bound form #status experimental <SIG>  
F;1-28/Domain: signal sequence #link MARS #status experimental <WATM>  
F;1-6/Domain: intracellular #status predicted <INT>  
F;7-28/Domain: transmembrane #status predicted <TMN>  
F;29-792/Domain: extracellular #status predicted <EXT>  
F;29-34/Domain: propeptide #link MARS #status experimental <PRO>  
F;35-792/Product: dipeptidyl-peptidase, soluble form #status experimental <WATS>  
F;83,90,148,217,227,319,521,686/Binding site: carboxylate (Asn) (covalent) #status predicted  
F;631/Active site: Ser #status experimental  
F;709,741/Active site: Asp, His #status predicted

Alignment Scores: 7.45e-13 Length: 792  
Pred. No.: 249.50 Matches: 122  
Score: 32.4% Conservative: 66  
Percent Similarity: 32.4% Mismatches: 188  
Best Local Similarity: 21.0% Indels: 205  
Query Match: 8.4% Gaps: 23  
DB: 1

US-10-825-632-6 (1-1669) x A39914 (1-792)

```
Qy 5 GGTACAGCAAACTCTAAAGTCATCTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGA 64
Db 258 GlyAlaValAsnProThrValLysPhePhele-----ValAsnThrAspSerLeuSer 275
Qy 65 AGGATCATAGATCATAGATAGGAACATAATCAACCTTTTGAGATTCATTGAAGGA 124
Db 276 SerThrThrThrProMetGlnThrAlaProAlaSerValThr---ThrGly 294
Qy 125 GTTGAATATATCCAGAGCTCGATGAGCTCTCTGAGGGAATAATGCTTGCTGCATCCTA 184
Db 295 AspHisTyrLeuCysAspValAlaTrpValSerGluAspArgIleSer----- 310
Qy 185 CTAGATCGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAATATTATATCCCA 244
Db 310 ----- 310
Qy 245 GTAGAAGATGATTTATGAGAAAGGACAGAGACTCATGATGAGTGCCTGATTCGTGACG 304
Db 311 -----LeuGlnTrpLeuArgArgIleGlnAsnTyr-----SerValMet 323
Qy 305 CCACATAATATCATGAGAAACACAGACATCTGATAAATATCATCATCATCTTTTAT 364
Db 324 AlaIleCysAspTyrAspLysThrAsnLeuValTrpAsnCysProThrThrGlnGluHis 343
Qy 365 GTT-----TTTCCCAAAAGTCAC 382
Db 344 IleGluThrSerAlaThrGlyTrpValGlyArgPheArgProAlaGluProHisPheThr 363
Qy 383 GAAGAGAAATGAGTTATTTTGGCTCTGAAATGCAAAACAGGTTCCGTCATTATATAC 442
Db 364 SerAspGlySerPheTyrLysIleValSerAspLysAspGlyTyrLysHis----- 381
Qy 443 AAATTTACATCTATTTAAGGAAACCAATATAACGATCCAGTGGGGCTGCCTGCT 502
Db 382 ---IleCysGlnPheGlnLysAspArgLys----- 390
Qy 503 CCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATATACCAGTGGTGAATGGAA 562
Db 391 -----ProGluGlnAspCysThrPheIleThrLysGlyAlaTrpGlu 404
Qy 563 GTTCTTGGCCGGCATGGATCATATATCCAAAGTTGATGAAGTCAGAAAGCTCGGTATATTT 622
Db 405 ValIle-----SerIleGluAlaLeuThrSerAspTyrLeuTyrTyrIle 419
Qy 623 GAAGGCAC---AAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAT 679
Db 420 SerAsnGluTyrLysGluMetProGlyGlyArgAsnLeuTyrLysIle----- 435
Qy 680 CTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTCCTTGCTGCATCAGT--- 736
```

RESULT 9  
CDHU26

```
Db 436 -----GlnLeuThrAsp-----HisThrAsnLysLysCysLeuSerCys 448
Qy 737 -----CAGCACTGTGCTCTCTTT-----ATAAGTAAGTATAGTAACACAG 775
Db 449 AspLeuAsnProGluArgCysGlnTyrTyrSerValSerLeuSerLysGluAlaLysTyr 468
Qy 776 AAGAAATCCACACTGT-----GTGTCCCTTTTCAAGCTATCAAGTCTCGAAGAT 823
Db 469 TyrGlnLeuGlyCysArgGlyProGlyLeuProLeuTyrThrLeuHisArgSerThrAsp 488
Qy 824 GACCAACTTGGCAAAACAAGAAATTTGGGCCACCATTTTGGATTCACAGAGTCCCTCTT 883
Db 489 GlnLysGluLeuArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----Leu 505
Qy 884 CTGACTATATCTCTCCAGAA-----ATTTCCTCTTTTGAAGTACTACTGGATTT 934
Db 506 GlnAspValGlnMetProSerLysLysLeuAspPheIleValLeuAsnGluThrArgPhe 525
Qy 935 ACATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACT 994
Db 526 ---TriPtyrGlnMetIleLeuProProHis---PheAspLysSerLysLysTyrProLeu 543
Qy 995 GTGCTGTTCATATATGCTGCTCT----- 1018
Db 544 LeuIleAspValTyrAlaGlyProCysSerGlnLysAlaAspAlaAlaPheArgPheAsn 563
Qy 1018 ----- 1018
Db 564 TrpAlaThrTyrLeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArgGly 583
Qy 1018 ----- 1018
Db 584 SerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsnLysArgLeuGlyThrLeuGlu 603
Qy 1018 ----- 1018
Db 604 ValGluAspGlnIleGluAlaAlaArgGlnPheLeuLysMetGlyPheValAspSerLys 623
Qy 1019 CAGTCTCTCTATTCGTGG----- 1036
Db 624 GlnValAlaIleTrpGlyTrpSerTyrGlyGlyTyrValThrSerMetValLeuGlySer 643
Qy 1037 -----GCCCCAGTCACTCTGTGATCTTCTAT 1063
Db 644 GlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGluTyrTyr 663
Qy 1064 GATACAGATACACGGAACGTTATATGGGT-----CACCTTGACCAAGTGAACAGGCG 1117
Db 664 AspSerValTyrThrGluArgTyrMetGlyLeuProThrProGluAspAsnLeuAspHis 683
Qy 1118 TATTACTTAGATCTGTGGCCATGCAAGCAAGAAAGTTCCTCTCTGAACCAAAATCGTTTA 1177
Db 684 TyrArgAsnSerThrValMetSerArgAlaGluAsnPhe-----LysGlnValGluTyr 701
Qy 1178 CTGCTCTTACATGTTTCTCGTAGAAGATGTCATTTTGGACATACACAGTATATTATCTG 1237
Db 702 LeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGlnGlnSerAlaGlnIleSer 721
Qy 1238 AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC 1297
Db 722 LysAlaLeuValAspAlaGlyValAspPheGlnAlaMetTyrTyrThrAspGluAspHis 741
Qy 1298 AGCATAAGAGTCTCTGAATCGGGAACATATTGAACATCTCTTTTGCACCTACCTCAA 1357
Db 742 GlyIleAlaSerSerThrAlaHisGlnHisIleTyrSerHisMetSerHisPheLeuGln 761
Qy 1358 GAA 1360
Db 762 Gln 762
```

dipeptidyl-peptidase IV (EC 3.4.14.5) - human  
 N:Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence revision 23-Aug-1996 #text change 09-Jul-2004  
 C:Accession: S24313; Y2408; J42408; B61136; S59510; I56154; S59857; S15520  
 R:Miiumi, Y.; Hayashi, Y.; Arakawa, F.; Ikenhara, Y.  
 Biochim. Biophys. Acta 1131, 333-336, 1992  
 A:Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a serine protease  
 A:Reference number: S24313; MUID:92329551; PMID:1352704  
 A:Accession: S24313  
 A:Molecule type: mRNA  
 A:Residues: 1-5, 'I', 8-766 <MIS>  
 A:Cross-references: UNIPROT:P27487; UNIPARC:UPI000016AB80; EMBL:X60708; NID:G35335; PIDN:R1DAMOL; D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, A.  
 J. Biol. Chem. 267, 4824-4833, 1992  
 A:Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer cell lines  
 A:Reference number: A42408; MUID:92165847; PMID:1347043  
 A:Accession: B42408  
 A:Molecule type: mRNA  
 A:Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-766 <DAR1>  
 A:Cross-references: UNIPARC:UPI0000052ACB; GB:M80536; NID:g181569; PIDN:AAA52308.1; PID:J181569  
 A:Experimental source: intestine  
 A:Note: this sequence corresponds with the author's translation  
 A:Accession: A42408  
 A:Molecule type: mRNA  
 A:Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-711, 'G', 713-766 <DAR2>  
 A:Cross-references: UNIPARC:UPI0000172A2B; GB:M80536; NID:g181569  
 A:Note: sequence extracted from NCBI backbone (NCBI:83986; NCBI:83988); this sequence is identical to the one in the GenBank database  
 R:Gorvel, J. P.; Ferrero, A.; Chambraud, L.; Rigal, A.; Bonicel, J.; Maroux, S.  
 Gastroenterology 101, 618-625, 1991  
 A:Title: Expression of aurocacin, a dipeptidyl-peptidase IV inhibitor, in human small intestine  
 A:Reference number: A61136; MUID:91317403; PMID:1677636  
 A:Accession: B61136  
 A:Molecule type: protein  
 A:Residues: 1-15, 'X', 17-22 <GOR>  
 A:Cross-references: UNIPARC:UPI0000172A2C  
 R:Boehm, S.K.; Gum, J.; J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S.  
 Biochem. J. 311, 835-843, 1995  
 A:Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a 5' flanking region  
 A:Reference number: S59510; MUID:96067599; PMID:7487939  
 A:Accession: S59510  
 A:Molecule type: DNA  
 A:Residues: 1-31 <BOE>  
 A:Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:g1195574; PIDN:AA835614.1; PID:R1TANAKA; T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg, H.  
 J. Immunol. 149, 481-486, 1992  
 A:Title: Cloning and functional expression of the T cell activation antigen CD26.  
 A:Reference number: I56154; MUID:92325476; PMID:1352530  
 A:Accession: I56154  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-436, 'S', 438-766 <TAN>  
 A:Cross-references: UNIPARC:UPI000004F7BF; GB:M74777; NID:g180082; PIDN:AAA51943.1; PID:R1ABBOTT; C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W.  
 Immunogenetics 40, 331-338, 1994  
 A:Title: Genomic organization, exact localization, and tissue expression of the human CD26 gene  
 A:Reference number: S59857; MUID:95012454; PMID:7927537  
 A:Accession: S59857  
 A:Molecule type: DNA  
 A:Residues: 1-436, 'S', 438-766 <ABB>  
 A:Cross-references: UNIPARC:UPI000004F7BF; EMBL:U13734  
 C:Genetics: GDB:DP44  
 A:Gene: GDB:DP44  
 A:Cross-references: GDB:125239; OMIM:102720  
 A:Map position: 2q24.3-2q24.3  
 A:Introns: 2/3; 32/1; 65/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/3  
 C:Superfamily: dipeptidyl-peptidase IV  
 C:Keywords: dipeptidyl-peptidase; glycoprotein; homodimer; proteinase; transmembrane protein  
 F:1-6/Domain: intracellular #status predicted <INT>  
 F:7-28/Domain: transmembrane #status predicted <TMN>  
 F:29-766/Domain: extracellular #status predicted <EXT>  
 F:85, 92, 150, 219, 229, 281, 321, 520, 685/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:630,708,740/Active site: Ser, Asp, His #status predicted

Alignment Scores:  
 Pred. No.: 5,82e-12 Length: 766  
 Score: 238.00 Matches: 120  
 Percent Similarity: 33.5% Conservative: 71  
 Best Local Similarity: 21.1% Mismatches: 193  
 Indels: 186  
 Query Match: 1 Gaps: 24  
 DB:

US-10-825-632-6 (1-1669) x CDH26 (1-766)

QY 5 GGTACAGCAAACTCTAAAGTCATCTTTAAAGATGTCAGAAATATGATGATGCTGAAGGA 64  
 |||||  
 Db 260 GlyAlaValAenProThrValLysPhePheVal-----ValAenThrAspSerLeuSer 277  
 |||||  
 QY 65 AGGATCATAGATGTCATAGATAGAACTAATCAACCTTTTGAGATCTCTATTGAAGGA 124  
 |||||  
 Db 278 SerValThrAenAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIle---Gly 296  
 |||||  
 QY 125 GTTGAATATATGCCAGAGCTGGATGCTCTCTAGGGGAAATATGCTTGGTCCATCCTA 184  
 |||||  
 Db 297 AspHisTyrLeuCysAspValThrPala----- 306  
 |||||  
 QY 185 CTAGATCGCTCCAGATCGCTACAGATAGTGTGATCTCACCTGAATATTATTATCCCA 244  
 |||||  
 Db 307 -----ThrGlnGluArgIleSerLeuGlnTrpLeu----- 316  
 |||||  
 QY 245 GTAGAGATGATGTTATGAAAGGAGAGACTCATGATCAGTCAGTCCTGATCTGTGACG 304  
 |||||  
 Db 317 -----ArgArgIleGlnAenTyr-----SerValMet 325  
 |||||  
 QY 305 CCACATAATTATCATGAAGAAACACAGACATCTCGATAAATATCATCATGACATCTTTCAT 364  
 |||||  
 Db 326 AspileCysAspTyrAspGluSerSerGlyArgTrp---AenCys----- 339  
 |||||  
 QY 365 GTTTTTCCTCCCAAGTCACAGAGAAATATGATTTATTTTTCCTCTGAATGCAAAACA 424  
 |||||  
 Db 340 LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrpValGlyArgPheArgPro 359  
 |||||  
 QY 425 GGTTTCCTGCTCAT-----TTATACAAATTTACATCTATTATTTAAAG 463  
 |||||  
 Db 360 SerGluProHisPheThrLeuAspGlyAenSerPheTyrLysIleSer-----Aen 377  
 |||||  
 QY 464 GAAAGCAAAATATAACGATCCAGTCAGTCGGCTGCTCCCAAGTATTTCAAGTGCCT 523  
 |||||  
 Db 378 GluGluGlyTyrArgHisIle-----CysTyrPheGlnIleAsp 390  
 |||||  
 QY 524 ATCAAGAGGAGATAGCAATACCAAGTGGTGAATGGAGAGTCTTGGCCGCGCATGGATCT 583  
 |||||  
 Db 391 LysLysAspCysThrPheIleThrLysGlyThrTrpGluValIleGly----- 406  
 |||||  
 QY 584 AATATCAAGTGTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACC---AAAGACTCC 640  
 |||||  
 Db 407 ---IleGluAlaLeuThrSerAspTyrTyrIleSerAenGluTyrLysGlyMet 425  
 |||||  
 QY 641 CTTTATAGATCACCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 700  
 |||||  
 Db 426 ProGlyGlyArgAenLeuTyrLysIleGlnLeuIleAspTyrThrLysValThrCysLeu 445  
 |||||  
 QY 701 ACTGACCGTGGCTACTCATCTTTGCTGCTAGT---CAGCACTGTGATCTCTTTATA 757  
 |||||  
 Db 446 -----SerCysGluLeuAenProGluArgCysGlnTyrTyrSer 458  
 |||||  
 QY 758 AGTAAGTATAGTAACCAAGAAATCCACAC-----TGT-----GTGTCC 796  
 |||||  
 Db 459 ValSerPheSerLysGluAlaLysTyrGlnLeuArgCysSerGlyProGlyLeuPro 478  
 |||||  
 QY 797 CTTTACAAGCTATCAAGTCTCTGAAGATGATGATGATGATGATGATGATGATGATGATG 856  
 |||||  
 Db 479 LeuTyrThrLeuHisSerSerValAenAspLysGlyLeuArgValLeuGluAsp---Aen 497  
 |||||  
 QY 857 ACCATTTTGGATTACAGAGGTCTCTCTCTGACTATATCTCTCTCCAGAA-----ATT 907

```
Db 498 SerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLysLysLeuAsp 515
Qy 908 TTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAGCCTCATGAT 967
Db 516 PheIleIleLeuAsnGluThrLysPhe---TriPyrGlnMetIleLeuProHis--- 533
Qy 968 CTACAGCCTGGAAGAAATATCTTACTGCTGCTCTTATATATATGTTGGTCTCT 1018
Db 534 PheAspLysSerLysLysPyrProLeuLeuLeuAspValLysAlaGlyProCysSerGln 553
Qy 1018 ----- 1018
Db 554 LysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIle 573
Qy 1018 ----- 1018
Db 574 IleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMethIle 593
Qy 1018 ----- 1018
Db 594 IleAsnArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe 613
Qy 1018 ----- 1018
Db 614 SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGly 633
Qy 1019 -----CAGTTGCTATTCTCTGGG 1036
Db 634 TyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaVal 653
Qy 1037 GCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTATATGGT--- 1093
Db 654 AlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArgTyrMetGlyLeu 673
Qy 1094 ---CACCTGACCAAGTACACAGGCTATTACTTAGGATCTGTGGCCATGCAACAGAA 1150
Db 674 ProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGlu 693
Qy 1151 AAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGTC 1210
Db 694 AsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAsnVal 711
Qy 1211 CATTTTGACATACCATATATTAATGAGTATTTTATGAGGCGCTGGAAGCCATATGAT 1270
Db 712 HisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGln 731
Qy 1271 TTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCCTGTAATCGGAGAACATAT 1330
Db 732 AlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIle 751
Qy 1331 GNACTGCACTTTTGGCACTACCTTCAAGAA 1360
Db 752 TyrThrHisMetSerHisPheIleLysGln 761
```

## RESULT 10

S23752  
dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse  
N:Alternate names: CD26 alpha subunit; THAM alpha subunit  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S23752; A46465; A56030  
R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.  
J. Biol. Chem. 267, 2200-2208, 1992  
A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di  
A:Reference number: S23752; MUID:92129288; PMID:1370813  
A:Accession: S23752  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-760 <MR>  
A:Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384  
R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;  
J. Immunol. 147, 447-454, 1991

A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase  
A:Reference number: A46465; MUID:91302787; PMID:11712807  
A:Accession: A46465  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <VIV>  
A:Cross-references: UNIPARC:UPI0000172A2E  
A:Experimental source: M4.T thymoma cells, Swiss nu/nu  
R:Note: sequence extracted from NCBI backbone (NCBIP:42236)  
R:Bernard, A.M.; Mattel, M.G.; Pierres, M.; Marguet, D.  
Biochemistry 33, 15204-15214, 1994  
A:Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.  
A:Reference number: A56030; MUID:95092780; PMID:7999781  
A:Accession: A56030  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 746-760 <BER>  
A:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620  
C:Genetics:  
A:Gene: CD26

C:Superfamily: dipeptidyl-peptidase IV  
C:Keywords: dipeptidyl-peptidase hydrolase; glycoprotein; transmembrane protein  
F:213,223,315,514,679/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:624,702,734/Active site: Ser, Asp, His #status predicted

Alignment Scores:  
Pred. No.: 2,8e-11 Length: 760  
Score: 230.00 Matches: 98  
Percent Similarity: 33.5% Conservative: 47  
Best Local Similarity: 22.6% Mismatches: 120  
Query Match: 7.8% Indels: 168  
DB: 1 Gaps: 19

US-10-825-632-6 (1-1669) x S23752 (1-760)

```
Qy 419 ARAACAGCTTTCCTGCTATTTATACAAAATACATCTATTTTAAAGGAACCAATATAA 478
Db 372 LysAspGlyTyrLysHisIleCysHisPhe----- 381
Qy 479 CGATCCAGTGGTGGGCTGCTCCAGTGATTTCAAG---TGTCCTATCAAGAGGAG 535
Db 382 -----ProLysAspLysLysAspCysThrPhe----- 390
Qy 536 ATAGCAATACAGTGGTGAATGGAAAGTTCTTGGCGGCATGGATCTAATATCAAGTT 595
Db 391 -----IleThrLysGlyAlaTrpGluValIle-----SerIleGluAla 403
Qy 596 GATGAAGTCAGAGCGTGGTATATTTTGAAGGCACC---NAAGACTCCCTCTTAGAGCAT 652
Db 404 LeuThrSerAspTyrLeuTyrTyrIleSerAsnGlnTyrLysGluMetProGlyGlyArg 423
Qy 653 CACCTGTAGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGTGACTGACCGTGGC 712
Db 424 AsnLeuTyrLysIle-----GlnLeuThrAsp----- 432
Qy 713 TACTCACATTTCTGCTGCATCAGT-----CAGCACTGTGCTCTTTTATA 757
Db 433 HisThrAsnValLysCysLeuSerCysAspLeuAsnProGluArgCysGlnTyrTyrAla 452
Qy 758 AGTAAGTATAGTAACCAAGAAATCCACAC-----TGT-----GTGTCC 796
Db 453 ValSerPheSerLysGluAlaLysTyrTyrGlnLeuGlyCysTrpGlyProGlyLeuPro 472
Qy 797 CTTTCAAGCTATCAAGTCCCTGGAAGATGACCAACTTGCACAAACAAAGGAATTTGGGCC 856
Db 473 LeuTyrThrLeuHisArgSerThrAspHisLysGluLeuArgValLeuGlu----- 489
Qy 857 ACCAATTTGGATTCAGCAGGT-----CCTCTTCCTGACTATATCTCTCCAGAA--- 904
Db 490 -----AspAsnSerAlaLeuAspArgMetLeuGlnAspValGlnMetProSerLys 506
Qy 905 -----ATTTCCTCTTTTGAAGTACTACTGGATTTCATTTGATGGGATGCTCTACAAG 958
```

Db	507	LysLeuaspPheIleValLeuAenGluThrArgPhe-----TrpTyrGlnMetIleLeuPro	525	Score: 226.00	Matches: 99		
Qy	959	CCTCATGATCTACAGCTCGAAGAAATATCTCTACTGTGCTGTTTCATATATGGTGGCTCT	1018	Percent Similarity: 34.3%	Conservative: 62		
Db	526	ProHis---PheAaspLysSerLysLysTyrProLeuLeuLeuAepValTyrAlaGlyPro	544	Best Local Similarity: 21.1%	Mismatches: 142		
Qy	1018	-----	1018	Query Match: 7.6%	Indels: 166		
Db	545	CysSerGlnLysAlaAaspAlaSerPheArgLeuAenTrpAlaThrTyrLeuAlaSerThr	564	DB: 2	Gaps: 18		
Qy	1018	-----	1018	US-10-825-632-6 (1-1669) x T25173 (1-779)			
Db	565	GluAenIleIleValAlaSerPheAepGlyArgGlySerGlyTyrGlnGlyApsLysIle	584	Qy	317	TATGAAGAACAACAGACATCTGGATAAAATATCCATGACATCTTTTCATGTTTCCCCAA	376
Qy	1018	-----	1018	Db	350	TyrLysTyrAlaSerLysArgTrpValThr---HisAaspPheHisSerIle---Thr	367
Db	585	MetHisAlaIleAenArgLeuGlyThrLeuGluValGluAepGlnIleGluAlaA	604	Qy	377	AGTCACGAAGAGAAATTTAGTTTATTTTTCCTCTGAAATGCAAAACAGAGTTTCCGTCAT	436
Qy	1019	-----	1036	Db	368	SerPheGluAepThrLeuPhePheLeuLeuPro-----His	379
Db	605	ArgGlnPheValLysMetGlyPheValAaspSerLysArgValAlaIleTrpGlyTrpSer	624	Qy	437	TTATACAAAATATACATCTATTATTAAGGAAAGCAAAATATAAACGATCCAGTCGTGGGTG	496
Qy	1036	-----	1036	Db	380	ApsLysArgAaspAenAlaPheGlnGlnValAlaSerLeuArgLeuSerHisGlyGlnLeu	399
Db	625	TyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGly	644	Qy	497	CCTGCTCCCAAGTGATTTTCAAGTGCTCTATCAAGAGGAGATAGCAATTACCAGTGGTGAA	556
Qy	1037	-----GCCCACTACTCTGGATCTCTATGATACAGATACAGCAAGCTTAT	1087	Db	400	ArgThrProLys-----PheLeuAenLeuGlyGlu	409
Db	645	IleAlaValAlaProValSerArgTrpGluTyrTyrAaspSerValTyrThrGluArgTyr	664	Qy	557	TGGGAAGTCTTGGCCGCATGATCTAATATCAAGTTGATCAAGTCAGAGGCTGTA	616
Qy	1088	ATGGGT-----CACCTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCCATG	1141	Db	410	TyrAaspValThrSerIleAenGlyIleAenLysGluThr-----ArgThrIle	425
Db	665	MetGlyLeuProIleProGluAaspAenLeuAepHisTyrArgAenSerThrValMetSer	684	Qy	617	TATTTTGAAGCCACAAAGATCCCTTTTAGACATCCTCTGACGTACGTACGTACGTACGTA	676
Qy	1142	CAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGAT	1201	Db	426	PhePheHisAlaAlaAlaProLysProSerHisArgSerLeuPheSerTyrSer-----	443
Db	685	ArgAlaGluHisPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAasp	702	Qy	677	AATCTGTGAGAGGTGACAAAGCTGACTGACCGTGGGTACTCACATTTCTTGTCTGCATCAGT	736
Qy	1202	GAGAATGTCCATTTTGACATACCACTATATATTACTAGTTTTTTAGTGGGCTGGAAAG	1261	Db	444	-----LeuAlaAaspGluSerArgAenSerAlaTyrCysIleSer	456
Db	703	AaspAenValHisPheGlnGlnSerAlaGlnIleSerLysValLeuValAaspAlaGlyVal	722	Qy	737	-----CAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACCAAGAAATCCACAC	787
Qy	1262	CCATATGATTTACAGATCTATCTCAGAGAGACACAGCATAAAGATTCTCTGAATCGGA	1321	Db	457	CysSerIleLysAenCysThrTrpAlaGlnAlaGlnMetAaspGlnMetLysThrAla	476
Db	723	AepPheGlnAlaMetTrpTyrThrAaspGluAepHisGlyIleAlaSerSerThrAlaHis	742	Qy	788	TGTGTGTCCTTTTCAAAAGCTATCAAGTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGT	823
Qy	1322	GAACATATGAACCTGATCTTTTGCACTACCTTCAAGAA	1360	Db	477	IleValSerCysLysGlyProAlaAlaProHisThrAlaIleValAenLeuThrArgMet	496
Db	743	GlnHisIleTyrSerHisMetSerHisPheLeuGlnGln	755	Qy	824	GACCCAACTTGCACAAACAAAG-----GAATTTTGGGCCACCAT	862
RESULT 11				Db	497	AepSerAepLysLysThrGluHisAlaAenLeuLeuTyrApsLysThrTyrGlnAenArg	516
T25173				Qy	863	TGGATTTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGT	922
hypothetical protein T23f1.7a - Caenorhabditis elegans				Db	517	ValGluGluAlaGly---LeuProValIleIleLysGluThrIle-----LysIle	532
C:Species: Caenorhabditis elegans				Qy	923	ACTACTGGATTATCATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCTCGGAAG	982
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004				Db	533	SerAaspAepPheAaspAlaLeuIleLysLeuSerIleProLysApsIleTyrAasnArgAap	552
C:Accession: T25173				Qy	983	AAATAT-----CCTACTGTGCTCTTTCATATATGCTGCTCT-----	1018
submitted to the EMBL Data Library, October 1996				Db	553	LysHisGlnAlaIleProLeuIleValHisValTyrGlyProAenAaspGlnAenThr	572
A:Accession: T25173				Qy	1018	-----	1018
A:Status: preliminary; translated from GB/EMBL/DBJ				Db	573	LysGluAlaThrGlnIleGlyIleGluGluValValAlaSerAlaSerGlnAlaAlaIle	592
A:Molecule type: DNA				Qy	1018	-----	1018
A:Residues: 1-779 <WIL>				Db	593	LeuArgIleAepGlyArgGlySerGlyGlyArgGlyTyrTrpLysTyrArgSerAlaIleTyr	612
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN:CAB03411.1;				Qy	1018	-----	1018
A:Experimental source: clone T23F1				Db	613	GlyGlnLeuGlyThrValGluValGluAaspGlnIleLysAlaIleLysValValLeuArg	632
C:Genetics:				Qy	1019	-----	1036
A:Map position: 5							
A:Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1							
C:Superfamily: dipeptidyl-peptidase IV							
Alignment Scores:							
Pred. No.:							





```
QY 1214 TTTCACATACCATATATTACTGAGTTTITTTAGTCAGGCTGGAAAGCCCATATGATTTA 1273
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 747 PheGlnAenSerAlaIleLeuIleAspGluLeuGlnAenArgGlyValAspPheAspLeu 766
QY 1274 CAGATCTATCTCTAGAGAGACACACATAAGAGTCTTGAATCGGAGAGACATATGAA 1333
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 767 MetValTyProAenGlnAlaHisSerLeuSerSerArgThrSer-----HisValVal 784
QY 1334 CTGCATCTTTTGCACTACCTTCAAGAA 1360
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 785 GlyLysMetThrHisPheLeuArgGln 793

RESULT 13
A:1793
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41793
R:Mada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wentholt, R.J.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
A:Title: Differential expression of two distinct forms of mRNA encoding members of a dip
A:Reference number: A41793; MUID:92108018; PMID:1729689
A:Accession: A41793
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-803 <WAD>
A:Cross-references: UNIPROT:P42659; UNIPARC:UPI00002A83C; GB:M76429; NID:g4087119; PIDN:
A:Note: sequence extracted from NCBI backbone (NCBI:P75136)
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase, glycoprotein
F:257,342/Binding site: carbohydrate (Aen) (covalent) #status predicted

Alignment Scores:
Pred. No.: 1,67e-10 Length: 803
Score: 221.00 Matches: 107
Percent Similarity: 32.8% Conservative: 65
Best Local Similarity: 20.4% Mismatches: 178
Query Match: 7.5% Indels: 174
DB: 2 Gaps: 19

US-10-825-632-6 (1-1669) x A41793 (1-803)
QY 131 TATATTGCCAGCTGGATGGACTCTCTGAGGAAATATGCT-----TGGTCCATCCTTA 184
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 320 TyrIleThrMetValLysTrpAlaThrSerThrLysValAlaValAenTrp----- 336
QY 185 CTAGATCGCTCCGAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATATCCCA 244
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 337 LeuSerArgAlaGlnAen-----ValSer----- 344
QY 245 GTAAAGATGATGTTATGAAAGCGAGAGACTATTGAGTCAGTCGCTGATTCTGTGACG 304
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 345 -----TATGAAGAAACAACAGACATCTGGATAAATATCCATGACATC 353
QY 305 CCACTAATTATC-----TATGAAGAAACAACAGACATCTGGATAAATATCCATGACATC 358
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 354 GlyValCysThrLysLysHisGluAspGluSerGluAlaTrpLeu----- 368
QY 359 TTTTCATGTTTTCCCAAGTCACCAAGAGGAATTCAGTTTATTTTTCCTCTGATGC 418
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 369 -----HisArgGlnAenGluProValPheSer----- 378
QY 419 AAAACAGGTTTCGTCATTATATACAAATTCATCTATTAAAG-----GAAACCAA 472
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 379 LysAspGly--ArgLysPhePheValArgAlaIleProGlnGlyGlnGlyLys 397
QY 473 TATAACAGATCAGTGGTGGGCTGCTCCTCAAGTGATTTCAAGTGCTCTATCAAAAGAG 532
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 398 PheTyRHisIleThrValSerSerSerGlnProAenSerSerAenAspAenIleGln--- 416
QY 533 GAGATAGCAATTCACAGTGGTGAATGGAGGTTCTTGGCGGATCGATCTAATATCCAA 592
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 417 -----SerIleThrSerGlyAspTrpAspValThr-----LysIleLeuSer 430
```

```
QY 593 GTTGATGAAGTCACAGGCTGGTATATTTTGAAGGCACCAAAAGACTCCCTTTAGAGCAT 652
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 431 TyrAspGluLysArgSerGlnIleTyrPheLeuSerThrGluAspLeuProArgArg 450
QY 653 CACCTGACGTAGTCAGTTTACGTAAATCTCTGGAGAGGTGACAAGG----- 697
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 451 GlnLeuTyrSerAlaSerThrVal-----GlySerPheAenArgGlnCysLeuSerCys 468
QY 698 ---CTGACTGACCGTGGCTACTCACAATTTCTTGCTGCATCAGTCAGCACTGTGACTCTTT 754
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 469 AspLeuValAspAenCysThrTyrPheSerAlaSerPheSerProGlyAlaAspPhePhe 488
QY 755 ATAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGT 814
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 489 LeuLeuLysCysGluGlyProGlyValPro---ThrValSerValHisAenThrThrAsp 507
QY 815 CTTCAAGATGACCACTTGCACAAACAAAGGAA---TTTGGGCCACCATTTTGGATTCA 871
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 508 LysLysLysMetPheAspLeuGluThrAenGluHisValGlnLysAlaIleSerAspArg 527
QY 872 GCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGA 931
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 528 GlnMetProLysValGluTyrArgLysIleGlu-----ThrAspAsp 541
QY 932 TTTACATGTTGGGATGCTCTACAAGCCTCATGATCTACAGCTGGGAAAGAAATATCTCT 991
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 542 TyrAsnLeuProIleGlnIleLeuLysProAlaThrPheThrAspThrAlaHisTyrPro 561
QY 992 ACTGTGCTGTCATATATATGCTGGTCTCT----- 1018
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 562 LeuLeuLeuValValAspGlyThrProGlySerGlnSerValAlaGluLysPheAlaVal 581
QY 1018 ----- 1018
Db 582 ThrTrpGluThrValMetValSerSerHisGlyAlaValValLysCysAspGlyArg 601
QY 1018 ----- 1018
Db 602 GlySerGlyPheGlnGlyThrArgLeuLeuHisGluValArgArgLeuGlySerLeu 621
QY 1018 ----- 1018
Db 622 GluGluLysAspGlnMetGluAlaValArgValMetLeuLysGluProTyrIleAspLys 641
QY 1019 ---CAGGTTGCTATTGCTGGG----- 1036
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 642 ThrArgValAlaValPheGlyLysAspTyrGlyGlyTyrLeuSerThrTyrLeuLeuPro 661
QY 1037 -----GCCCCAGTCACT 1048
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 662 AlaLysGlyAspGlyGlnAlaProValPheSerCysGlySerAlaLeuSerProIleThr 681
QY 1049 CTGTGGATCTTCTATGATACAGATACACGAGAGCTTATATGGCTACCCCTGACCAAT 1108
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 682 AspPheLysLeuTyrAlaSerAlaPheSerGluArgTyrLeuGlyLeuHisGlyLeuAsp 701
QY 1109 GAACAGGGTATATCTAGGATCTGTGGCATGACGAGCAAGAAAGTCCCTCTGAAACCA 1168
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 702 AsnArgAlaTyrGluMetAlaLysValAlaHisArgValSerAlaLeu-----GluGly 719
QY 1169 AATCGTTTACTGCTCTTATCATGGTTTCTGATGAGAAATGTCATTTTGCACATPACAGT 1228
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 720 GlnGlnPheLeuValIleHisAlaThrAlaAspGluLysIleHisPheGlnHisThrAla 739
QY 1229 ATATTACTAGTTTTTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAG 1288
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 740 GluLeuIleThrGlnLeuIleLysGlyLysAlaAenTyrSerLeuGlnIleTyrProAsp 759
QY 1289 GAGAGACACAGCATAAAGAGTTCTCGAATCGGGAGAACATATATGAACTCGCATCTTTTGCAC 1348
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 760 GluSerHisTyrPheSerSerAlaAlaLeuGlnGlnHisLeuHisArgSerIleLeuGly 779
```

```
QY 1349 TACCTTCAGAA 1360
Db 780 PhePheValGlu 783
RESULT 14
168600
dipeptidyl aminopeptidase like protein - human
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
R;Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum Mol Genet. 2, 1037-1039, 1993
A;Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
A;Reference number: 154331; MUID:93372805; PMID:810397
A;Accession: I68600
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-803 <RES>
A;Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:g306707; PIDN:AAA35761.1; PID:
C;Superfamily: dipeptidyl-peptidase IV
Alignment Scores:
Pred. No.: 4,45e-10 Length: 803
Score: 216.00 Matches: 111
Percent Similarity: 32.7% Conservative: 79
Best Local Similarity: 19.1% Mismatches: 179
Query Match: 7.3% Indels: 212
DB: 2 Gaps: 23
US-10-825-632-6 (1-1669) x I68600 (1-803)
QY 5 GGTACAGCAATCTTAAGTCACTTTTAAGTGTGAGAAATGATGATGCTGAAGGA 64
Db 286 GlySerGluAsnProSerLeuHis-----ValIleGlyLeuAsnGly 301
QY 65 AGGATCATAGTCTCATAGTAAGGAATCAATTTCAACCTTTTGAGATTCTATTGGAAGGA 124
Db 302 ProThrHisAspLeu-----GluMetMetProProAspAspProArgMetArgGlu 318
QY 125 GTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATATGCT-----TGTTCC 178
Db 319 Tyr---TyrIleThrMetValLysTrpAlaThrSerThrLysValAlaValThrTrp--- 336
QY 179 ATCTCATAGATCGCTCCAGATCCGCTACAGATAGTGTGATCTACCTGATATTATTT 238
Db 337 -----LeuAsnArgAlaGlnAsn-----ValSer----- 344
QY 239 ATCCAGTAGNAGATGATGTTATGGAAGGAGCAGAGACTCATTGAGTCACTGCTGCTGATTTCT 298
Db 345 -----IleLeuThrLeuCysAspAla 351
QY 299 GTGACGCCACTAATATTC-----TATGAAGAAACACACAGACATCTGGATAAATATCCAT 352
Db 352 ThrThrGlyValCysThrLysLysHisGluAspGluSerGluAlaTrpLeuHis----- 369
QY 353 GACATCTTTCATGTTTTCCTCCAAAGTCACGAGAGAGAAAT----- 394
Db 370 -----ArgGlnAsnGluGluProValPheSerLysAspGlyArg 382
QY 395 GAGTTTATTTTGGCTCTGATGCAAAACAGGTTTCCGT-----CATTTATACAAATTAACA 451
Db 383 LysPhePheIleArgAlaIleProGlnGlyGlyArgGlyLysPheThrHisIleThr 402
QY 452 TCTATTTTAAAGGAAGAAATATAAACGATCAGTGGTGGGCTGCTGCTCCCAAGTGAAT 511
Db 403 -----ValSerSerSerGlnProAsnSerSerAsnAspAsnIleGln----- 416
QY 512 TTCAGTGTCTCTATCAAAGAGGAGATAGCAATTTACAGTGTGGTGAATGGAGTCTTGGC 571
Db 417 -----SerIleThrSerGlyAspTrpAspValThr--- 426
QY 572 CGGATGGATCTAATATCAAGTGTGATGAGTCAAGGCTGATATATTGGAAGGCACC 631
Db 763 TyrPheThrSerSerSerLeuLysGlnHisLeuTyrArgSerIleIleAsnPheVal 782
```

```
Db 427 -----LysIleLeuAlaTyrAspGluLysGlyAsnLysIleTyrPheLeuSerThr 443
QY 632 AAAGACTCCCTTTAGAGCATCACCTGTACCTAGTACGTACGTAAATCTCGAGAGGTG 691
Db 444 GluAspLeuProArgArgGlnLeuTyrSerAlaAsn----- 456
QY 692 ACAAGCTGACTGACCGTGGCTACTACAT-----TCTTCTGCATCAGTCAGCAC 742
Db 457 -----ThrGluGlyAsnPheAsnArgGlnCysLeuSerCysAspLeuValGluAsn 473
QY 743 TGTGACTTCTTTATAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTCCCTTAC 802
Db 474 CysThrTyrPheSerAlaSerPheSerHisSerMetAsp-----PhePheLeuLeu 490
QY 803 AAGCTATCAAGTCT-----GAAGATGACCCCAACTTGCAGCAAAAG 844
Db 491 LysCysGluGlyProGlyValProMetValThrValHisAsnThrThrAspLysLysLys 510
QY 845 GAATTT-----TGGGCCACCATTTTGGATTTCAGCAGGTCT 880
Db 511 MetPheAspLeuGluThrAsnGluHisValLysLysAlaIleAsnAspArgGlnMetPro 530
QY 881 CTTCCTGACTATACTCTCCAGAAATTTCTCTTTTGAAGTACTTCTGGATTTCACATTG 940
Db 531 LysValGluTyrArgAspIleGluIle-----AspAspTyrAsnLeu 544
QY 941 TATGGGATGCTCTACAGGCTCATGATCTACACGCTGGAAAGAAATATCTTACTGTGCTG 1000
Db 545 ProMetGlnIleLeuLysProAlaThrPheThrAspThrHisTyrProLeuLeuLeu 564
QY 1001 TTCATATPATGCTGGTCT----- 1018
Db 565 ValValAspGlyThrProGlySerGlnSerValAlaGluLysPheGluValSerTrpGlu 584
QY 1018 ----- 1018
Db 585 ThrValMetValSerSerHisGlyAlaValValLysCysAspGlyArgGlySerGly 604
QY 1018 ----- 1018
Db 605 PheGlnGlyThrLysLeuLeuHisGluValAlaArgArgLeuGlyLeuGluGluLys 624
QY 1019 -----CAGTT 1024
Db 625 AspGlnMetGluAlaValArgThrMetMetLeuLysGluGlnTyrIleAspArgThrArgVal 644
QY 1025 GCTATTGCTGG----- 1036
Db 645 AlaValPheGlyLysAspTyrGlyTyrLeuSerThrTyrIleLeuProAlaLysGly 664
QY 1037 -----GCCCCAGTCACTCTGTGGATC 1057
Db 665 GluAsnGlnGlyGlnThrPheThrCysGlySerAlaLeuSerProIleThrAspPheLys 684
QY 1058 TTCTATGATACAGGATACACGGAACGTTATATGGTCACTCCCTGACCAAGATGAACAGGC 1117
Db 685 LeuTyrAlaSerAlaPheSerGluArgTyrLeuGlyLeuHisGlyLeuAspAsnArgAla 704
QY 1118 TATTACTTAGATCTGTGGCATGCAAGCAAGAAAGTCTCCCTCTCAACCAATCGTTA 1177
Db 705 TyrGluMetThrLysValAlaHisArgValSerAlaLeu-----GluGlnGlnPhe 722
QY 1178 CTGCTCTTACATGTTTCTCTGGATGAGAATGCCATTTTGCATACATACCATATATTCTG 1237
Db 723 LeuIleIleHisProThrAlaAspGluLysIleHisPheGlnHisThrAlaGluLeuIle 742
QY 1238 AGTTTATAGTCCAGGCTGGAAAGCCCATATGATTTACAGATCTATCTCTCAGAGAGACAC 1297
Db 743 ThrGlnLeuIleArgGlyLysAlaAsnTyrSerLeuGlnIleTyrProAspGluSerHis 762
QY 1298 ACATAAGAGTTCCTGAATCGGGAACATATGAACATGCACTCTTTTGCACCTTCAA 1357
Db 763 TyrPheThrSerSerSerLeuLysGlnHisLeuTyrArgSerIleIleAsnPheVal 782
```



Db 825 TyrPheThrSerSerLeuLysGlnHisLeuTyrArgSerIleIleAsnPheVal 844

Qy 1358 GAA 1360

|||

Db 845 Glu 845

Search completed: May 2, 2006, 01:37:46  
Job time : 89.6336 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 00:42:23 ; Search time 76.878 Seconds  
(without alignments)  
3063.366 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 2957  
Sequence: 1 aacagtcagcaaatccta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEl=frame+ n2p.model -DEV=xlp  
-Q=/abss/ABSSWEB.spool/US10825632/runat\_01052006\_105946\_3226/app\_query.fasta-1  
-DB=UniProt -QMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=opto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p  
-USER=US10825632.OCGN 1\_1\_694 @runat\_01052006\_105946\_3226 -NCFU=6 -ICPU=3  
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2422	81.9	898	1 DPP8 HUMAN	Q6v1x1 homo sapien
2	2326	78.7	892	1 DPP8_MOUSE	Q80ya7 mus musculu
3	1494.5	50.5	863	1 DPP9 HUMAN	Q86t12 homo sapien
4	1488.5	50.3	862	1 DPP9_MOUSE	Q8bvg4 mus musculu
5	1462	49.4	923	2 Q4SBM6 TETNG	Q4sbm6 tetraodon n
6	1389	47.0	847	2 Q6R222_XENLA	Q6gr22 xenopus lae
7	1187.5	40.2	508	2 Q75273_HUMAN	Q75273 homo sapien
8	894	30.2	621	2 Q7PTT8_ANOGA	Q7pct8 anopheles g
9	889	30.1	886	2 Q7QBK1_ANOGA	Q7qbki anopheles g
10	867.5	29.3	740	2 Q5TTK8_ANOGA	Q5ttk8 anopheles g
11	841.5	28.5	1053	2 Q9VC20_DROME	Q9vc20 drosophila
12	841.5	28.5	1113	2 Q9VCL9_DROME	Q9vc19 drosophila
13	718.5	24.3	557	2 Q5TXJ2_ANOGA	Q5txj2 anopheles g
14	613.5	20.7	803	2 Q54U01_DICDI	Q54u01 dictyosteli
15	462	15.6	927	2 Q965K3_CABEL	Q965k3 caenorhabdi
16	462	15.6	931	2 Q44987_CABEL	Q44987 caenorhabdi

17	449	15.2	552	2	Q8GUJ7_ARATH	Q8guj7 arabidopsis
18	449	15.2	746	2	Q9FNF6_ARATH	Q9fnf6 arabidopsis
19	444	15.0	596	2	Q6K880_ORYSA	Q6k880 oryza sativ
20	429.5	14.5	738	2	Q9ASE0_CAUCR	Q9ase0 caulobacter
21	415.5	14.1	763	2	Q8EAB7_SHEON	Q8eab7 shewanella
22	413	14.0	895	2	Q61CU7_CABBR	Q61cu7 caenorhabdi
23	391	13.2	743	2	Q5QX36_IDILO	Q5qx36 idiomarina
24	386	13.1	238	2	Q5TYG2_ANOGA	Q5tyg2 anopheles g
25	383	13.0	720	2	Q5NMW8_ZYMMO	Q5nmw8 zymomonas m
26	379.5	12.8	745	2	Q6FJ17_PSED	Q6fj17 pseudomonas
27	369.5	12.5	736	2	Q8A028_BACTN	Q8a028 bacteroides
28	359.5	12.2	736	2	Q5LGU5_BACFN	Q5lgu5 bacteroides
29	359.5	12.2	766	2	Q4TNP1_9SPHN	Q4tnp1 erythrobact
30	356.5	12.1	736	2	Q64XP9_BACFR	Q64xp9 bacteroides
31	355.5	12.0	751	2	Q4UFD3_XANCP	Q4upd3 xanthomonas
32	355.5	12.0	751	2	Q8PJV8_XANCP	Q8pjv8 xanthomonas
33	355	12.0	741	2	P95782_XANNA	P95782 xanthomonas
34	351	11.9	745	2	Q5H5W8_XANOR	Q5h5w8 xanthomonas
35	346	11.7	757	2	Q8PFD7_XANAC	Q8pfd7 xanthomonas
36	322	10.9	769	2	Q5QVY7_IDILO	Q5qv77 idiomarina
37	320.5	10.8	850	2	Q6H9E3_9TRYP	Q6h9e3 trypanosoma
38	318	10.8	947	2	Q4WX13_ASPFU	Q4wx13 aspergillus
39	312.5	10.6	874	2	Q7SI80_EMENI	Q7si80 emericella
40	312.5	10.6	880	2	Q5B934_EMENI	Q5b934 aspergillus
41	312	10.6	901	2	Q96VT7_ASPNG	Q96vt7 aspergillus
42	309.5	10.5	771	2	Q42B12_ASPNG	Q42b12 aspergillus
43	308.5	10.4	765	2	Q4WPH9_ASPFU	Q4wph9 aspergillus
44	308.5	10.4	765	2	O14425_ASPFU	O14425 aspergillus
45	306.5	10.4	778	2	Q512P7_MAGGR	Q512p7 magnaporthe

#### ALIGNMENTS

RESULT 1  
DPP8\_HUMAN STANDARD; PRT; 898 AA.  
ID AC Q6V1X1; Q7Z4C8; Q7Z4E3; Q7Z4E4; Q8HBM5; Q9NXP4;  
AC Q9HBM3; Q9HBM4; Q9HBM5; Q9NXP4;  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII) (DP8)  
DE (prolyl dipeptidase DPP8) (Dipeptidyl peptidase IV-related protein 1)  
DE (DPP8-1)  
GN Name=DPP8; Synonyms=DPP81; ORFNames=MSTP097, MSTP135, MSTP141;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF 334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION, CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION, RP AND SUBCELLULAR LOCATION.  
RC TISSUE=Placenta;  
RX MEDLINE=20467194; PubMed=11012666;  
RA Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W., RA Correll M.D.;  
RT "Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homology, DPP8.";  
RL Eur. J. Biochem. 267:6140-6150(2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME REGULATION, BIOPHYSICO-CHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND RP TISSUE SPECIFICITY.  
RC TISSUE=Testis;  
RX PubMed=12662155; DOI=10.1042/BJ20021914;  
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akinsanya K.O.;  
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member of an emerging subgroup of serine proteases.";  
RL Biochem. J. 373:179-189(2003).



RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).  
 RP TISSUE=Testis;  
 RA Sha J.H., Zhou Z.M., Li J.M.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4), AND NUCLEOTIDE  
 RP SEQUENCE [LARGE SCALE MRNA] OF 211-898 (ISOFORM 2).  
 RC TISSUE=Hepatoma, and Placenta;  
 RX PubMed:14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,  
 RA Nagahashi K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,  
 RA Shiratori A., Sudo H., Hosioki T., Kaku Y., Kodaira H., Kikkawa E.,  
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,  
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa K.,  
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,  
 RA Fujimori K., Tanai H., Kimata S., Watanabe M., Hiraoka S., Chiba Y.,  
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,  
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
 RA Nomura Y., Togihara S., Komai F., Hara R., Takeuchi K., Arita M.,  
 RA Imose N., Musashino K., Yuuki F., Oshima A., Sasaki N., Aotsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Suzuki H.,  
 RA Hishigaki H., Watanabe K., Kumagai A., Itakura S., Fukukami B.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y., Okamoto S.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani K., Kawajima T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsuura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togaishi T., Oyama H., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45(2004).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).  
 RC TISSUE=Testis;  
 RX MEDLINE=2388257; PubMed:12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.H.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalón D.K., Muzny D.M., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,  
 RA Richards S., Worley K.C., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Raheij J., Holton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitling M., Madan A., Young A.C., Green E.D., Dickson M.C.,  
 RA Blakeley R.W., Touchman A.J., Schmutz J., Myers R.M.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Skalek U., Smailus D.B.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.B.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 561-898.  
 RC TISSUE=Aorta;  
 RA Zhao B., Xu H.S., Tong Y.K., Sheng H., Qin B.M., Liu Y.Q., Liu B.,  
 RA Wang X.Y., Zhang Q., Song L., Gao Y., Zhang C.L., Ye J., Ji X.J.,  
 RA Liu B.H., Lu H., Chen J.Z., Cai M.Q., Zheng W.Y., Teng C.Y., Liu Q.,  
 RA Yu L.T., Lin J., Gong Q., Zhang A.M., Gao R.L., Hui R.T.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP MUTAGENESIS OF GLU-275; SER-755; ASP-833 AND HIS-865, CATALYTIC  
 RP ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, AND SUBCELLULAR LOCATION.  
 RX PubMed:12534281; DOI=10.1021/bi026946g;  
 RA Ajami K., Abbott C.A., Obradovic M., Gysbers V., Kaehne T.,  
 RA McCaughan G.W., Gorrell M.D.;  
 RT "Structural requirements for catalysis, expression, and dimerization  
 RT in the CD26/DPPIV gene family.";  
 RL Biochemistry 42:694-701(2003).  
 RN [8]  
 RP BIOPHYSICOCHEMICAL PROPERTIES.  
 RX PubMed:15039077; DOI=10.1016/j.pep.2003.12.019;  
 RA Chen Y.-S., Chien C.-H., Goparaju C.M., Hsu J.T.-A., Liang P.-H.,  
 RA Chen X.;  
 RT "Purification and characterization of human prolyl dipeptidase DPP8 in  
 RT Sf9 insect cells.";  
 RL Protein Expr. Purif. 35:142-146(2004).  
 RN [9]  
 RP ENZYME REGULATION.  
 RX PubMed:15664838; DOI=10.1016/j.bmc.2004.11.023;  
 RA Jiaang W.-P., Chen Y.-S., Hsu T., Wu S.-H., Chien C.-H., Chang C.-N.,  
 RA Chang S.-P., Lee S.-J., Chen X.;  
 RT "Novel isoindoline compounds for potent and selective inhibition of  
 RT prolyl dipeptidase DPP8.";  
 RL Bioorg. Med. Chem. Lett. 15:687-691(2005).  
 CC -1- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal  
 CC dipeptides from proteins having a Pro or Ala residue at position  
 CC 2. May play a role in -cell activation and immune function.  
 CC -2- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-  
 CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided  
 CC Zaa is neither Pro nor hydroxyproline.  
 CC -3- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine  
 CC proteinase inhibitor 4-(2-aminoethyl)benzenesulfonyl fluoride  
 CC (ABEPS), and by di-isopropylfluorophosphate. Specifically inhibited  
 CC by isoindoline derivatives.  
 CC -4- BIOPHYSICOCHEMICAL PROPERTIES:  
 CC Kinetic parameters:  
 CC KM=208 uM for Ala-Pro-AMC;  
 CC KM=130 uM for Ala-Pro-AFC;  
 CC KM=120 uM for H-Ala-Pro-pNa;  
 CC KM=1420 uM for H-Ala-Ala-pNa;  
 CC KM=310 uM for H-Arg-Pro-pNa;  
 CC KM=2050 uM for H-Asp-Pro-pNa;  
 CC KM=480 uM for H-Gly-Pro-pNa;  
 CC pH dependence:  
 CC Optimum pH is 7.4-8.5. Little activity below pH 6.5;  
 CC -5- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -6- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=6;  
 CC Name=1:  
 CC IsoId=06VIX1-1; Sequence=Displayed;  
 CC Name=2:  
 CC IsoId=06VIX1-2; Sequence=VSP\_013864;  
 CC Name=3:  
 CC IsoId=06VIX1-3; Sequence=VSP\_013860;  
 CC Name=4:  
 CC IsoId=06VIX1-4; Sequence=VSP\_013860, VSP\_013862;  
 CC Name=5:  
 CC IsoId=06VIX1-5; Sequence=VSP\_013863;  
 CC Name=6:  
 CC IsoId=06VIX1-6; Sequence=VSP\_013861;  
 CC -7- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in  
 CC testis, placenta, prostate, muscle and brain.  
 CC -8- INDUCTION: In activated T-cells.  
 CC -9- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.  
 CC -10- CAUTION: Ref.4 (BAA91059) sequence differs from that shown due to  
 CC a frameshift in position 486.  
 CC -11- CAUTION: Ref.6 (AAQ13650 and AAQ13623) sequences differ from that  
 CC shown due to several frameshifts.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not  
CC removed.

CC	EMBL; AF221634; AAG29766.1; -; mRNA.
DR	EMBL; AF221635; AAG29767.1; -; mRNA.
DR	EMBL; AF221636; AAG29768.1; -; mRNA.
DR	EMBL; AF221637; AAG29769.1; -; mRNA.
DR	EMBL; AY172659; AAQ17261.1; -; mRNA.
DR	EMBL; AY354203; AAQ63887.1; -; mRNA.
DR	EMBL; AK000290; BAA91059.1; ALT FRAME; mRNA.
DR	EMBL; AK027826; BAB55395.1; ALT INIT; mRNA.
DR	EMBL; BC030688; RAH30688.3; -; mRNA.
DR	EMBL; BC040203; RAH40203.1; ALT INIT; mRNA.
DR	EMBL; AF176739; RAQ13657.1; ALT INIT; mRNA.
DR	EMBL; AF175225; RAQ13650.1; ALT FRAME; mRNA.
DR	EMBL; AF173382; RAQ13623.1; ALT FRAME; mRNA.

Alignment Scores:					
Pred. No.:	2,49e-195	Length:	898		
Score:	2422.00	Matches:	455		
Percent Similarity:	82.3%	Conservative:	0		
Best Local Similarity:	82.3%	Mismatches:	0		
Query Match:	81.9%	Indels:	100		
DB:	1	Gaps:	1		

UUS-10-825-632-6 (1-1669) x DPP8 HUMAN (1-898)

Qy	2	ACAGGTACACGAAATCCTTAAGTGCATCTTTTAAGATGTCAGAAATATGATTGATGCTGAA	61
Db	334	ThrGlyThrAlaasnProLysVaIThrPheLysMetSerGluIleMetIleAspAlaGlu	353
Qy	62	GGNAGGATCATAGATGTCATAGATAAGAACTAATTCAAACCTTTTCAGATTCTTATTGAA	121
Db	354	GlyArgIleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGlu	373
Qy	122	GGAGTTGAATATATTCCAGAGCTGGATGGACTCCTCAGGGAAAAATATGCTTGGTCCATC	181
Db	374	GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle	393
Qy	182	CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCATCCTGAAATTTATTATC	241
Db	394	LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle	413
Qy	242	CCAGTTAAGATGATGCTTATGAAAGGCAGAGACTCATTGAGTCAGTGCTGATCTGTG	301
Db	414	ProValGluAspAspValMetGluArgGlnArgIleGluSerValProAspSerVal	433
Qy	302	ACGCCACTAATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTT	361
Db	434	ThrProLeuIleIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe	453
Qy	362	CA TGTTTTCCTCCAAAGTCACGAAGAGAAATATGAGTTATTTTGGCTCTGAAATCAAA	421
Db	454	HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys	473
Qy	422	ACAGGTTCCGTCATTATACAAATTCATCTATTTTAAAGGAACCAATATAACGA	481
Db	474	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	493
Qy	482	TCCAGTGGTGGGCTGCTGCTCCAAAGTGATTTTCAAGTGTCTCATCAAAGAGGAGATGCA	541
Db	494	SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla	513
Qy	542	ATTACCAAGTCGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAA	601
Db	514	IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu	533
Qy	602	GTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTAC	661
Db	534	ValArgargLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr	553
Qy	662	GTAGTCAGTTTACGTAATCTCTGAGAGGTGACAAGGCTGACTGACCGGTGCTACTCACAT	721

RESULT 2  
DPP8 MOUSE  
ID\_DPP8 MOUSE  
STANDARD;  
PRT; 892 AA.

PRT; 892 AA.



Db 428 ThrProLeuIleIleTyrGluGluThrThrAspIleIleTrpIleAsnIleHisAspIlePhe 447  
Qy 362 CATGTTTTCCTCCCAAGTACAGAGAGAAATTTAGTTTATTTTGGCTCTCTGAATGCAAA 421  
Db 448 HisValPheProGlnThrHisGluAspGluIleGluPheIlePheAlaSerGluCysLys 467  
Qy 422 ACAGTTTCCTCATTTATACAAAATTTACATCTATTTTAAAGAAAGCAATATAACGA 481  
Db 468 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 487  
Qy 482 TCCAGTGGTGGCTGCTCTCCAGTGATTTCAAGTCTCTATCAAGAGGAGATAGCA 541  
Db 488 SerSerGlyLeuProIleProSerAspPheLysCysProIleLysGluGluIleThr 507  
Qy 542 ATTACAGTGGTGAATGGGAAGTTCTTGGCCGGCATGATCTAATATCAAGTTGTATGAA 601  
Db 508 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleTrpValAspGlu 527  
Qy 602 CTCAGAGGCTGTATATTTGAGGACCAAGACTCCCTTTAGAGCATCACTGTAC 661  
Db 528 AlaArgLysLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyr 547  
Qy 662 GTAGTCAGTTACGTAATCTCGAGAGGTGACAGCGTCACTGACCGTGGCTACTACAT 721  
Db 548 ValThrSerTyrAlaAsnProGlyGluValValArgLeuThrAspArgGlyTyrSerHis 567  
Qy 722 TCTGTGTCATCAGTCAGCATGTCGACTCTTTTATAGTAAATGATAGTAAACGAGAAT 781  
Db 568 SerCysLysSerArgHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 587  
Qy 782 CCACACTGTGTCCCTTACAGCTATCAAGTCTCAAGTCTCAAGTACCCCACTGCAAAACA 841  
Db 588 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProValHisLysThr 607  
Qy 842 AAGGAATTTTGGCCACCATTTTGGATTGAGTTCAGCAGTCTCTCTCTGACTATCTCTCCA 901  
Db 608 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 627  
Qy 902 GAAATTTCTTTTGAAGTACTACTGGATTACATTTGATGGATGCTCTCAAGCCT 961  
Db 628 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 647  
Qy 962 CATGATCTACAGCTCGAAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCAG 1021  
Db 648 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 667  
Qy 1021 ----- 1021  
Db 668 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 687  
Qy 1021 ----- 1021  
Db 688 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 707  
Qy 1021 ----- 1021  
Db 708 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGly 727  
Qy 1021 ----- 1021  
Db 728 LeuGlnTyrLeuAlaSerGlnTyrAspPheIleAspLeuAspArgValGlyIleHisGly 747  
Qy 1021 ----- 1021  
Db 748 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 767  
Qy 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACACGAA 1081  
Db 768 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheThrAspThrGlyTyrThrGlu 787  
Qy 1082 CGTTATATGGTCACTGACAGAAATGACAGGCTATTACTAGATCTGTGGCATG 1141

Db 788 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 807  
Qy 1142 CAACGAGAAAGTTCCTCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGAT 1201  
Db 808 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 827  
Qy 1202 GAGBAATCCATTTTCACATACCACATATATTACTAGTGTATTTTGTAGTGGGCTGAAAG 1261  
Db 828 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 847  
Qy 1262 CCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTTCTCGAATCGGA 1321  
Db 848 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 867  
Qy 1322 GAACTATGAACTGACATCTTTTGGACTACTCTCAAGAAACCTTGGATCAGTATTCCT 1381  
Db 868 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 887  
Qy 1382 GCTCTAAAGTGATA 1396  
Db 888 AlaLeuLysValIle 892  
RESULT 3  
DPP9 HUMAN  
ID DPP9 HUMAN STANDARD; PRT; 863 AA.  
AC Q86T12; Q6AL10; Q6ZMT2; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WDX8;  
AC Q96NT8; Q9BVE3;  
DT 13-SEP-2005 (Rel. 48, Created)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DE Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)  
DE (Dipeptidyl peptidase-like protein 9) (DPLP9) (Dipeptidyl peptidase  
IV-related protein 2) (DPRP-2).  
GN Name=DPP9; Synonyms=DPRP2;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.  
RX MEDLINE=22347328; PubMed=12459266; DOI=10.1016/S0378-1119(02)01059-4;  
RA Olsen C., Wagtmann N.;  
RT "Identification and characterization of human DPP9, a novel homologue  
of dipeptidyl peptidase IV.";  
RL Gene 299:185-193(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), CATALYTIC ACTIVITY,  
BIOPHYSICO-CHEMICAL PROPERTIES, ENZYME REGULATION, TISSUE SPECIFICITY,  
RP AND SUBCELLULAR LOCATION.  
RC TISSUE=Colon;  
RX PubMed=12662155; DOI=10.1042/BJ20021914;  
RA Qi S.Y., Riviere P.J., Trojanar J., Junien J.-L., Akinsanya K.O.;  
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member  
of an emerging subgroup of serine proteases.";  
RL Biochem. J. 373:179-189(2003).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), PARTIAL NUCLEOTIDE SEQUENCE  
[MRNA] (ISOFORM 2), CATALYTIC ACTIVITY, BIOPHYSICO-CHEMICAL PROPERTIES,  
RP TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.  
RX PubMed=15245913; DOI=10.1016/j.bbexp.2004.03.010;  
RA Ajami K., Abbott C.A., McCaughan G.W., Gorrell M.D.;  
RT "Dipeptidyl peptidase 9 has two forms, a broad tissue distribution,  
cytoplasmic localization and DPV-like peptidase activity.";  
RL Biochim. Biophys. Acta 1679:18-28(2004).  
RN [4]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
RC TISSUE=Placenta; and Skin;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Kzywinaki M.I., Skalek U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[5]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 30-863 (ISOFORM 3),  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 272-863 (ISOFORM 2), AND  
RP PARTIAL NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 5).  
RC TISSUE=Glial tumor, Ovary, Spleen, and Trachea;  
RX PubMed-14702033; DOI=10.1038/ng1285;  
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,  
RA Nagatani K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,  
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,  
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,  
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa K.,  
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,  
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,  
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,  
RA Kusano Y., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
RA Nomura Y., Togai S., Komai F., Hara R., Takeuchi K., Arita M.,  
RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,  
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
RA Hishigaki H., Watanabe T., Sugiyama A., Itakura S., Watanabe M.,  
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ikeda Y., Okamoto S.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki S., Itoh T., Shigeta K., Senba T.,  
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Mizushima-Sugano J., Sato H., Sato H., Tanigami Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
RT "Complete sequencing and characterization of 21,243 full-length human  
RT cDNAs";  
RL Nat. Genet. 36:40-45 (2004).  
[6]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 209-863 (ISOFORM 4), AND  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 298-863 (ISOFORM 2).  
RC TISSUE=Melanoma;  
RG The German cDNA consortium;  
Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal  
CC dipeptides from proteins having a pro or Ala residue at position  
CC 2.  
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-  
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided  
CC Zaa is neither Pro nor hydroxyproline.  
CC -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-  
CC (2-aminoethyl)benzenesulphonyl fluoride (AEBSF), and by di-  
CC isopropyl fluorophosphate.  
CC -!- BIOPHYSICOCHEMICAL PROPERTIES:  
CC Kinetic parameters:  
CC KM=161 uM for Ala-Pro-AMC;  
CC KM=180 uM for Ala-Pro-AFC;  
CC pH dependence:  
CC

CC Optimum pH is 7.5-8.5. Little activity below pH 6.5;  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; cytosol.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=5;  
CC Name=1; Synonyms=Short; Sequence=Displayed;  
CC IsoId=Q86T12-1; Sequence=VSP\_013865;  
CC Name=2; Synonyms=Long; Sequence=VSP\_013865;  
CC Note=Incomplete sequence;  
CC Name=3;  
CC IsoId=Q86T12-3; Sequence=VSP\_013867, VSP\_013868;  
CC Name=4;  
CC IsoId=Q86T12-4; Sequence=VSP\_013869;  
CC Name=5;  
CC IsoId=Q86T12-5; Sequence=VSP\_013865, VSP\_013866;  
CC Note=Incomplete sequence;  
CC TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in  
CC liver, heart and muscle, and lowest levels in brain.  
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.  
CC -!- CAUTION: Ref.6 (CAD39039) sequence differs from that shown due to  
CC frameshifts in positions 432 and 460.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC  
CC -----  
CC EMBL: AF452102; AAL47179.1; -- mRNA.  
CC EMBL: AY172460; AAO17262.1; -- mRNA.  
CC EMBL: AF542510; AAO73880.2; -- mRNA.  
CC EMBL: AV374518; AAO31119.1; -- mRNA.  
CC EMBL: BC000970; AAH09370.1; -- mRNA.  
CC EMBL: BC037848; AAH37948.1; -- mRNA.  
CC EMBL: AK054656; BAB70784.1; ALT\_INIT; mRNA.  
CC EMBL: AK131000; BAC1362.1; -- mRNA.  
CC EMBL: AK131499; BAC8150.1; -- mRNA.  
CC EMBL: AL034376; CAD39039.3; ALT\_FRAME; mRNA.  
CC EMBL: CR627380; CAH10477.1; -- mRNA.  
CC MEROPS: S09.019; --  
CC Ensembl: ENSG00000142002; Homo sapiens.  
CC HGNC: HGNC:18648; DPP9.  
CC MIM: 608258; --  
CC InterPro: IPR001375; Peptidase S9.  
CC InterPro: IPR002469; Peptidase S9B.  
CC InterPro: IPR000379; Ser. esters.  
CC Pfam: PF00930; DPPIV\_N; 1.  
CC Pfam: PF00326; Peptidase S9; 1.  
CC Alternative splicing; Aminopeptidase; Hydrolase; Protease;  
CC Serine protease.  
CC ACT\_SITE 730 730 Charge relay system (By similarity).  
CC ACT\_SITE 808 808 Charge relay system (By similarity).  
CC ACT\_SITE 840 840 Charge relay system (By similarity).  
CC VARSPPLIC 1 1 M -> LRRVPCVRCRPLPLPGSQSRMSRDEAPLD  
CC PGRPAQSRGPRPTSVSHACSWNGSLDLEGTALLRSAAE  
CC RLMEKVKLRDLKENTGSRFSINSEGAER (in  
CC isoform 2 and isoform 5).  
CC /FTid=VSP\_013865.  
CC Missing (in isoform 5).  
CC /FTid=VSP\_013866.  
CC OLVANSFKGIYLRNLNLASLGAV -> SAHLPPPPPHH  
CC PPNSSPPLACQL (in isoform 3).  
CC /FTid=VSP\_013867.  
CC Missing (in isoform 3).  
CC /FTid=VSP\_013868.  
CC Missing (in isoform 4).  
CC /FTid=VSP\_013869.  
CC I -> N (in Ref. 3).  
CC C -> W (in Ref. 5; BAC8150).  
CC L -> P (in Ref. 5; BAC18643).  
CC G -> C (in Ref. 5; BAB70784).  
CC SEQUENCE 863 AA; 98263 MW; 40F80B78E26CDE5 CRC64;  
SQ

Alignment Scores:			
Pred. No.:	5,19e-117	Length:	863
Score:	1494.50	Matches:	281
Percent Similarity:	62.9%	Conservative:	69
Best Local Similarity:	50.5%	Mismatches:	105
Query Match:	1	Indels:	101
DB:		Gaps:	2
US-10-825-632-6 (1-1669) x DPP9_HUMAN (1-863)			
QY	2	ACAGGTACACAAATCTTAAGTCACTTTTAAAGATGTGTCAGAAATATGATGCTGAA	61
DB	308	ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln	327
QY	62	GGAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGGAGATCTCTATTGAA	121
DB	328	GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro	347
QY	122	GGAGTTGAATATATTGCCAGAGCTGGATGACTCTCTGAGGGAAATATGCTTGTGTCATC	181
DB	348	LybValGluTyrlleAlaArgAlaGlyTrpThrArgAspGlyLysTyrlleAlaTipAlaMet	367
QY	182	CTATAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTATC	241
DB	368	PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIle	387
QY	242	CCAGTAGAAGATGTTATGGAAAGCGAGAGACTCATTGAGTCAGTSCCTGATTCTGTG	301
DB	388	ProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal	407
QY	302	ACGCCACTAATTATCTATGAAGAAACAAACAGACATCTCGATATAATATCCATGACATCTTT	361
DB	408	GlnProTyrlleValValTyrlleGluValThrAsnValTrpIleAsnValHisAspIlePhe	427
QY	362	CAHTTTTTCCCAAGTCAAC---GAAGAGAAATTTGATTTATTTTTGCTCTCGAATGC	418
DB	428	TyrProPheProGlnSerGluGlyGluAspGluLeuCyPheLeuArgAlaAsnGluCys	447
QY	419	AAACAGCTTTCCTGATTTATACAAATTTACATCTATTATTAAGAGAAAGCAATATAAA	478
DB	448	LysThrGlyPheCysHisLeuTyrlleValThrAlaValLeuLysSerGlnGlyTyrlleAsp	467
QY	479	CGATCCAGTGTGGCTGCTGCTCCCAAGTGATTTCAAGTGTCCTATCAAAAGAGAGATA	538
DB	468	TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGluIle	487
QY	539	GCAATTACCATGTGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCAAGTTGAT	598
DB	488	AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAsn	507
QY	599	GAAGTCAGAGGCTGTATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTG	658
DB	508	GluGluThrLysLeuValTyrllePheGlnGlyThrLysAspThrProLeuGluHisIleu	527
QY	659	TACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAAGGCTCACTGACCGTGGCTACTCA	718
DB	528	TyrValValSerTyrlleAlaAlaGlyGluIleValArgLeuThrProGlyPheSer	547
QY	719	CATTCTGTGTCATCAGTCAGCATGTGATCTTTTATAAGTAAGTATAGTAACACAGAG	778
DB	548	HisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrlleSerSerValSer	567
QY	779	AAATCCACATGTGTGCTCCCTTACAGCTATCAAGTCTCAAGTCTGAAGATGACCCCACTGCAAA	838
DB	568	ThrProProCysValHisValTyrlleLysLeuSerGlyProAspAspProLeuHisLys	587
QY	839	ACAAAGGAATTTTGGCGCACCATTTTGGATTACAGCAGGTCCTCTTCTGACTATACCTCT	898
DB	598	GlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrlleValPro	607
QY	899	CCAGAAATTTTCTTTTGAAGAGTACTACTGAGTATTACATCTGAGTATTGATGAGTCTACACAG	958

Db	608	ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrlleGlyMetIleTyrllys	627
QY	959	CCTCATGATCTACAGCCTCGAAGAAATATCTACTGTCTCTTTCATATATGTTGGTCT	1018
Db	628	ProHisAlaLeuGlnProGlyLysHisProThrValLeuPheValTyrlleGlyPro	647
QY	1019	CAG-----	1021
Db	648	GlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrlleArgLeuAsnThrLeu	667
QY	1021	-----	1021
Db	668	AlaSerLeuGlyTyrlleAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeu	687
QY	1021	-----	1021
Db	688	ArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGlu	707
QY	1021	-----	1021
Db	708	GlyLeuGlnPheValAlaGluLysTyrlleGlyPheIleAspLeuSerArgValAlaIleHis	727
QY	1021	-----	1021
Db	728	GlyTrpSerTyrlleGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe	747
QY	1022	---GTTGCTATTGCTGGGGCCCGAGTCACTCTGTGTGATCTTCTATCATACAGGATACAG	1078
Db	748	LysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrlleAspThrGlyTyrlle	767
QY	1079	GAAGCTTATATGGTCACTCCCTGACAGATGAACAGGGCTATATCTAGGATCTGGCC	1138
Db	768	GluArgTyrlleMetAspValProGluAsnAsnGlnHisGlyTyrlleGluAlaGlySerValAla	787
QY	1139	ATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTTCTTACATGTTCTCTG	1198
Db	788	LeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleHisGlyPheLeu	807
QY	1199	GATGAGATGTCATTTTGCACATACAGTATATATCTAGTGTGTTTGTAGTGGCTGGA	1258
Db	808	AspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGly	827
QY	1259	AAGCATATGATTTACAGATCTATCTCAGGAGAGACAGACAGATAGAGTCTCTCAATCG	1318
Db	828	LysProTyrlleGlnLeuGlnIleTyrlleProAsnGluArgHisSerIleArgCysProGluSer	847
QY	1319	GGAGAACATATGAACTGTCATCTTTGCACTACCTTCAAGAAACCTT	1366
Db	848	GlyGluHisTyrlleGluValThrLeuLeuHisPheLeuGlnGluTyrlleu	863
RESULT 4			
DPP9_MOUSE			
ID	DPP9_MOUSE	STANDARD;	PRT; 862 AA.
AC	Q8BYG4; Q8KAM9; Q8BWT9;		
DT	13-SEP-2005 (Rel. 48, Created)		
DT	13-SEP-2005 (Rel. 48, Last sequence update)		
DT	13-SEP-2005 (Rel. 48, Last annotation update)		
DE	Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)		
DE	(Dipeptidyl peptidase-like protein 9) (DPLP9).		
GN	Name=Dpp9;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridea; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]_TaxID=10090;		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).		
RC	STRAIN=C57BL/6J; TISSUE=Liver, and Olfactory bulb;		
RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;		
RA	Kazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,		
RA	Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,		
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,		
RA	Baldarelli R., Hill D.F., Bult C., Hume D.A., Quackenbush J.,		







```
QY 302 ACGCACTAATATCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 361
Db      |||
QY 407 GlnProPheValIleTyrGluValThrAsnValTrpIleAsnValHisAspIlePhe 426
Db      |||
QY 362 CATGTTTTCCTCCCAAGTCAAC---GAAGAGAAATAGAGTTATTTTGGCTCTGAATGC 418
Db      |||
QY 427 HisProPheProGlnAlaGluGlyGlnGlnAspPheCysPheLeuArgAlaAsnGluCys 446
QY 419 AAAACAGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAATATAAA 478
Db      |||
QY 447 LysThrGlyPheCysHisLeuTyrArgValThrValGluLeuLysThrLysAspTyrAsp 466
QY 479 CGATCCAGTGGTGGCTCGCTCCCAAGTATTTCAGTGCCTATCAAGAGAGAGATA 538
Db      |||
QY 467 TrpThrGluProLeuSerProThrGluAspGluPheLysCysProIleLysGluVal 486
QY 539 GCAATTACCACTGGTGAATGGAGATTCTTGGCGGCATGGATCTAATATCCAAATTCAT 598
Db      |||
QY 487 AlaLeuThrSerGlyGluTrpGluValLeuSerArgHisGlySerLysIleTrpValAsn 506
QY 599 GAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGATCAACCTG 658
Db      |||
QY 507 GluGlnThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLysLeu 526
QY 659 TAAGTAGTCAGTTACGTAATTCCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCA 718
Db      |||
QY 527 TyrValValSerTyrGluSerAlaGlyGluIleValArgLeuThrThrLeuGlyPheSer 546
QY 719 CATCTTCTGCATCAGTCAGCTGACTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAG 778
Db      |||
QY 547 HisSerCysSerMetSerGlnSerPheAspMetPheValSerHisTyrSerSerValSer 566
QY 779 AATCCACACTGTGTCCCTTACAGCTATCAAGTCCTGAAGATGACCCAACTTGCAAA 838
Db      |||
QY 567 ThrProProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys 586
QY 839 ACAAGGAATTTGGCGCACCATTTGGATTACAGAGTCTCTCTCTGACTATACTCTCT 898
Db      |||
QY 587 GlnProArgPheTrpAlaSerMetMetGluAlaAlaAsnCysProProAspTyrValPro 606
QY 899 CCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAG 958
Db      |||
QY 959 CCTCATGATCTACAGCTGGAAGAAATATCTCTGCTGCTGTTCATATATAGTGGTCTCT 1018
Db      |||
QY 627 ProHisThrLeuGlnProGlyArgLysHisProThrValLeuPheValTyrGlyGlyPro 646
QY 1019 CAG----- 1021
QY 647 GlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeu 666
QY 1021 ----- 1021
QY 667 AlaSerLeuGlyTyrAlaValValIleAspGlyArgLysSerCysGlnArgGlyLeu 686
QY 1021 ----- 1021
QY 687 HisPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGlu 706
QY 1021 ----- 1021
QY 707 GlyLeuGlnTyrValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis 726
QY 1021 ----- 1021
QY 727 GlyTrpSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe 746
QY 1022 ---GTTGCTATTGCTGGGGCCCGAGTCACTCTCTGGAATCTTCTATGATACAGGATACAG 1078
Db      |||
QY 747 LysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 766
QY 1079 GAACGTTATATGGGTCAACCTGACCAGAAATGAACAGGCTATTTACTTAGGATCTGGGCC 1138
```

```
Db      |||
QY 767 GluArgTyrMetAspValProGluAsnAsnGlnGlnGlyTyrGluAlaGlySerValAla 786
QY 1139 ATGCAAGCAGAAAAATGTTCCCTCTGACACAAATCGTTACTGCTCTTACTGTTCTCTG 1198
Db      |||
QY 787 LeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleHisGlyPheLeu 806
QY 1199 GATGAGATGTCACATTTTGCACATACCAGTATATTTACTGAGTTTTCTAGTGGGCTGGA 1258
Db      |||
QY 807 AspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGly 826
QY 1259 AAGCCATATGATTTTACAGATCTATCTCTCAGGAGACAGACAGCATTAAGATCTCTGATCG 1318
Db      |||
QY 827 LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysArgGluSer 846
QY 1319 GGAGAACATTAATGAACATGTCATCTTTTGCATCTCTTCAAGAAACCTT 1366
Db      |||
QY 847 GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluHisLeu 862

RESULT 5
Q4SBM6_TETNG
ID Q4SBM6_TETNG PRELIMINARY; PRT; 923 AA.
AC Q4SBM6;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Chromosome 15 SCAF14667, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00020903001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Salfan D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Athouart V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Bionnet C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J., Bosak S.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAEE01014667; CAG01956.1; -; Genomic_DNA.
FT NON TER 1 1
FT NON TER 923 923
SQ SEQUENCE 923 AA; 105211 MW; 1397023B2004D009 CRC64;

Alignment Scores:
Pred. No.: 2.94e-114 Length: 923
Score: 1462.00 Matches: 286
Percent Similarity: 56.5% Conservative: 67
Best Local Similarity: 45.8% Mismatches: 102
Query Match: 49.4% Indels: 170
DB: 2 Gaps: 6

US-10-825-632-6 (1-1669) x Q4SBM6_TETNG (1-923)
```





```

Db      732 LysValAlaIleAlaGlyAlaProValThrLeuTyrMetAlaTyrAspThrGlyTyrThr 751
QY      1079 GAACGTTATATGGTCAACCTGACAGAAATGACAGGGCTATTACTTAGGATCTGTGGCC 1138
Db      752 GluArgTyrMetGluThrProGluThrAsnGlnMetGlyTyrGluAlaGlySerAlaAla 771
QY      1139 ATCAACGACGAAAGTCCCTCTGACCAAAATCGTTTACTCTCTTACATGTTTCTGTG 1198
Db      772 LeuGlnValAspLysLeuProAsnGluProTyrArgLeuLeuIleLeuHisGlyPheLeu 791
QY      1199 GATCAGATGTCCTCAATTTGGCATACACAGTATATTAATCTAGAGTTTCTTGTAGCGGCTGGA 1258
Db      792 AspGluAsnValHisPhePheHisThrAsnPheLeuLeuSerGlnLeuIleArgAlaGly 811
QY      1259 AAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGACATGAAGAGTTCTCGAATCG 1318
Db      812 LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGluSer 831
QY      1319 GGAGAACATATTAAGACTGCACTTTTGGCACTACCTCTCAAGAAACCTT 1366
Db      832 GlyGluHisTyrGluIleThrLeuLeuHisPheLeuGlnGluHisLeu 847

RESULT 7
O75273 HUMAN
ID O75273 HUMAN PRELIMINARY; PRT; 508 AA.
AC O75273;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE R26984.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schulz K., Gordon L., Regala A., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do D., Regala W., Terry A., Barnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankheim M., Anico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC005594; AAC33801.1; -; Genomic_DNA.
DR MEROPS; S09.019; -.
DR Ensembl; ENSG00000142002; Homo sapiens.
DR GO; GO:0003824; F.catalytic activity; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER
FT 1
SQ SEQUENCE 508 AA; 57750 MW; 2F4D6645BE2D2C89 CRC64;

Alignment Scores!
Pred. No.: 3,75e-91 Length: 508
Score: 1187.50 Matches: 241
Percent Similarity: 56.2% Conservative: 64
Best Local Similarity: 44.4% Mismatches: 93
Query Match: 40.2% Indels: 145
DB: 2 Gaps: 7

US-10-825-632-6 (1-1669) x O75273_HUMAN (1-508)
QY 68 ATCATAGATGTCATAGATAAGCACTAATTCACCTTTTGTAGATTCTATTGAGGAGTT 127
Db 1 IleValSerThrGlnGluGlySerGlnValGlnProPheSerSerLeuPheProLysVal 20
QY 128 GAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAAATATGCTTGGTCCATCTACTA 187

```

```

Db      21 GluTyrIleAlaIleArgAlaGly-----AlaTrpAlaMetPheLeu 33
QY      188 GATCGCTCCAGACTCCGCTACAGATAGTGTTCATCTCACCTGAATATTATTATCCAGTA 247
Db      34 AspArgProGlnGlnTrpLeuGlnLeuValLeuProAlaLeuPheIleProSer 53
QY      248 GAAGATGATGTATTGAAAGGACGAGACTCAATGAGTCAGTCGCTGATCTCTGTGACGCA 307
Db      54 ThrGluAsnGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnPro 73
QY      308 CTAATTATCTATGAAGAAACACAGACATCTGATAAATATCATCATCTTTCATGTT 367
Db      74 TyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrPro 93
QY      368 TTTCCCAAGTACAC---GAAAGAGAAATTTAGTTTATTTTTCCTCTGAAATGCAAAACA 424
Db      94 PheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThr 113
QY      425 GGTTCCTCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAACGATCC 484
Db      114 GlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSer 133
QY      485 AGTGTGGCTGCTCTCCCAAGTGATTTCAAGTGCCTATCAAGAGAGAGATACAATT 544
Db      134 GluProPheSerProGly-----GluGlyGluGlnSerLeu 145
QY      545 ACCAGTCGTGAATGGGAGTTCTTGGCCGCGATGATCTAATATCAAGTTGATGAATC 604
Db      146 ThrAsnAlaIleTrp-----ValAsnGluGlu 154
QY      605 AGAAGCTGTATATTTTGAAGCACCAAGACTCCCTTTAGACATCACTGTACGTA 664
Db      155 ThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeuTyrVal 174
QY      665 GTCAATTACGTAATCTCGAGAGGTGACACAGCTGACGACCGTGGTACTCATCTCT 724
Db      175 ValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSer 194
QY      725 TGCTGCATAGTCAGCACCTGTGACTTCTTTTAAAGTAAGTATAGTAACACAGAGAATCCA 784
Db      195 CysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrPro 214
QY      785 CACTGTGTGTCCTTTTACCAAGCTTCAAGTCTCTGAGATGACCCCAACTTGCAAAACAAG 844
Db      215 ProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLysGlnPro 234
QY      845 GAATTTTGGGCCACCATTTTGGATTTCAGAGGTCTCTTCTCTGACTATATCTCTCCAGAA 904
Db      235 ArgPheTrpAlaSerMetMetGluAlaIle-----Lys 245
QY      905 ATTTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGGATGCTCTCAAGCCTCAT 964
Db      246 IlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHis 265
QY      965 GATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCAG--- 1021
Db      266 AlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnVal 285
QY      1021 ----- 1021
Db      286 GlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSer 305
QY      1021 ----- 1021
Db      306 LeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPhe 325
QY      1021 ----- 1021
Db      326 GluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeu 345
QY      1021 ----- 1021

```

```

Db 346 GlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrp 365
QY 1022 -----GTT 1024
Db 366 SerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysVal 385
QY 1025 GCTATTGTCGGGGCCCGCTACTCTGTGGATCTTCTATGATACAGGATACAGGAACGT 1084
Db 386 AlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArg 405
QY 1085 TATATGGTGCACCTCCAGCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCCAA 1144
Db 406 TyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHis 425
QY 1145 GCAGAAAGTTCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTTCTGGATGAG 1204
Db 426 ValGluLysLeuProAsnGluProAsnArgLeuLeuIleHisGlyPheLeuAspGlu 445
QY 1205 AATGTCATTTTGCATACACAGTATATTAATGAGTCTTTTGTAGGAGGCTGGAAAGCCA 1264
Db 446 AsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysPro 465
QY 1265 TATGATTTA-----CAGATCTATCTCTCAGGAGACAC 1297
Db 466 TyrGlnLeuGlnValAlaLeuProValSerProGlnIleTyrProAsnGluArgHis 485
QY 1298 AGCATAGAGTTCCTGATCGGAGACATTAATGAGTGCATCTTTTGCACCTTCA 1357
Db 486 SerIleArgCysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGln 505
QY 1358 GAAACCTT 1366
Db 506 GlutyrLeu 508

RESULT 8
ID Q7PTT8_ANOGA PRELIMINARY; PRT; 621 AA.
AC Q7PTT8;
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE ENSANGP0000015447 (Fragment).
GN ORFNames=ENSANG00000012958;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST.
RC "Anopheles gambiae Sequence Committee;
RT Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST.
RC The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008287; EAA03335.3; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DB NON_TER 1

```

```

FT NON TER 621
SQ SEQUENCE 621 AA; 70546 MW; F6EA8463A343BBF3 CRC64;

Alignment Scores:
Pred. No.: 2,42e-66 Length: 621
Score: 894.00 Matches: 211
Percent Similarity: 46.1% Conservative: 69
Best Local Similarity: 34.8% Mismatches: 147
Query Match: 30.2% Indels: 180
DB: 2 Gaps: 14

US-10-825-632-6 (1-1669) x Q7PTT8_ANOGA (1-621)
QY 5 GTTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATGAATGATTGATGCTGAAGGA 64
Db 41 GlyThrProAsnAlaLysSerLysLeuLysLeuValGlnPheArgLeuSerGluAsnLeu 60
QY 65 AGGATCATAGATCATAGATAAGGAACCTTAATCAACCTTTTGAGATTCTATTTTGAAGGA 124
Db 61 ArgIleThrAspValCysIleLysGluLeuGlnCysProLeuThrPheAlaPheProTrp 80
QY 125 GTTGAATATATTTGCCAGAGCTGGATGAGCTCTGAGGAAATAATATGCTGTGTCATCTTA 184
Db 81 LeuGluTyrIleValArgValGlyTrpThrProAspSerArgTyrValTrpAlaGlnLeu 100
QY 185 CTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACCTGAATATTATTATCCCA 244
Db 101 LeuAspArgProGlnGlnArgLeuGluLeuValLeu-----LeuPro 114
QY 245 GTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTGCCT----- 292
Db 115 Val---AspAsnPheCysGluIleTyrSerSerSerSerSerLeuProProArgLysSer 133
QY 293 -----GATTCTGTGACGCCACTA---ATTATCTAT 319
Db 134 SerSerGlyTrpArgSerProLeuAspLysSerThrThrArgProLeuGlnValIleTyr 153
QY 320 GAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGT 379
Db 154 ThrGluThrSerSerSerTrpValAsnValHisAspValLeuGln---PheValGluLeu 172
QY 380 CACGAAGAGAAATGAGTTTATTTTTCCTCTGATGCAAAACAGGTTTCGTCATTTA 439
Db 173 SerGluGlnGluValThrPheLeuTrpAlaSerGlu---GluSerGlyPheArgHisLeu 191
QY 440 TACAAAATTACATCTATTTTAAAGAAAGCAATATAACGATCCAGTGGTGGGCTGCCT 499
Db 192 TyrLeuValThrSerSerLeuSerProNanGluVal-----SerSerGlyGlyAlaGly 209
QY 500 GCTCCAAGTGTATTTCAAGTGTCTCT----- 523
Db 210 Ala---ThrAspHisSerLeuProSerMetAlaCysIleGlySerThrLeuValAlaArg 228
QY 524 ATCAAGAGAGGATAGCAATTACAGTGTGAATGGAAAGTTCTTTGGCGGCGATGATCT 583
Db 229 IleValGlnLysValThrLeuThrGlyGlyAspTrpGluValLeuGlyArg----- 245
QY 584 AATATCCAAGTTGATGAAGTCAGAGCTGGTATATATTTTGAAGGCGCAACAGACTCCCT 643
Db 246 AsnValTrpTyrAspArgValArgGlnLeuValTyrPheMetGlyLeuArgGluThrPro 265
QY 644 TTAGAGCATCTCTGTAGTACGTACGTACGTAAATCTCGAGAGGTCAGACAGGCTGACT 703
Db 266 LeuGluLysHisLeuTyrValSerLeuAlaGlnProAsnGlnLeuArgLeuLeuThr 285
QY 704 GACCGTGGCTACTCACATCTTCTGTCATCAGTCAGCTGTCATCTTTTATAAGTAAG 763
Db 286 MetProGlyTyrSerPheThrValGluPheAsnAspAspCysThrLeuPheLeuGlnThr 305
QY 764 TATAGTAACAGAGAAGATCCACACTGTGTGTCCTTTTACAAGCTA-----TCAAGT 814
Db 306 TyrCysAsnIleSerThrLeuProSerTrpGluLeuValArgIleAlaHisAspSerAsn 325

```







188 GlyThrProAsnAlaLysSerLysLeuValGlnPheArgLeuSerGluAsnLeu 207  
QY 65 AGGATCATAGTGTATAGTAAGAACTAACTTTCAGATTCTATTGGAAGGA 124  
Db ArgIleThrAspValCysLeuGlnCysProLeuThrPheAlaPheProTyr 227  
QY 125 GTTGAATATATGCCAGAGCTGGATCGACTCTCGAGGAAATAATATGCTTGCCTCATCTA 184  
Db LeuGluTyrIleValArgValGlyTyrThrProAspSerArgTyrValTyrPheAlaGlnLeu 247  
QY 185 CTAGATCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATATTATATCCCA 244  
Db LeuAspArgProGlnArgLeuGluValLeuLeuPheValAspAsnPheCysGlu 267  
QY 245 GTAGAAGATGATGTTATGAAAGGCAGAGACTCATTGAGTCACTGCTCT----- 292  
Db IleTyrSerSer-----SerSerArgProProAlaGlyAla 279  
QY 293 -----GATTCTGTGAGCCACTA---ATTATCTATGAAGAAACAACAGACATCTGG 340  
Db HisArgLysSerThrThrArgProLeuGlnValIleTyrThrGluThrSerSerTyr 299  
QY 341 ATAAATATCCATCACATCTTTCATGTTTTCCTCCAAAGTCACGAGAGGAAATATGAGTTT 400  
Db ValAsnValHisAspValLeuGln---PheValGluLeuSerGluGlnValThrPhe 318  
QY 401 ATTTTTCCTCTCAATGCAAAACAGGTTTCCTGCTATTTATATAAATATACATCTATTTTA 460  
Db LeuTrpAlaSerGlu---GluSerGlyPheArgHisLeuTyrLeuValThrSer----- 335  
QY 461 AAGGAAGCAAAATATAACAGATCCAGTGTGGGCTGCTCCAGTGAATTCMAAGTGT 520  
Db LysValSerSerGlyGlyValGlyAla---ThrAspHisSerLeu 349  
QY 521 CCT-----ATCAAGAGAGGATAGCAATT 544  
Db ProSerMetAlaCysIleGlySerThrLeuValAlaArgIleValGlnLysValThrLeu 369  
QY 545 ACCAGTGTGTAAGTGGAAAGTCTTTCGCGGCATGATCTAATATCCAAAGTGTGTAAGTC 604  
Db ThrGlyGlyAspTrpGluValLeuGlyArg-----AsnValTyrTyrAspArgAla 386  
QY 605 AGAAGGTGTATATTTTGAAGCCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTA 664  
Db ArgGlnLeuValTyrPheMetGlyLeuArgGluThrProLeuGluLysHisLeuTyrVal 406  
QY 665 GTCAAGTTACGTAATCTCGAGAGTGACACAGGCTGACTGACCGTGGCTACTCACATTCT 724  
Db ValSerLeuAlaGlnProAsnGlnLeuArgLeuLeuThrMetProGlyTyrSer----- 424  
QY 725 TGCTGCATCAGTCAGCACTGTCACCTTTTATAAGTAAATAGTAACT----- 772  
Db -----PheThrValGluPheAsnAspValSerValPhe 435  
QY 773 CAGAAGAATCCACACTGTGTCCCTTTTACAGACTATCAAGTCTCGAAGATGACCCCACT 832  
Db GluArgLysGlyAlaCysProAsnThrAlaAsnGlyAsnGlyCysSerHisGlyProThr 455  
QY 833 TGCNAACAAGAAATTTGGGCCACCATTTGGATTTCAGCAGGCTCTCTCTCGACTAT 892  
Db ProProProThrProIleAspAlaLeuArgLeuCysSerValGlyTyrLeuThrGluGly 475  
QY 893 ACTGCT-----CCAGAAATTTCTCTTTTGAAGTACTACTCGGA 931  
Db GlyProSerGluAsnThrGlnTyrAsnProSerIleHisSerProGlnIleSerSerGly 495  
QY 932 TTTACATGATGATGGATGCTTACAGGCTCATGATCTACAGCTGGAAAGAAATATCTCT 991  
Db AspValLeuTyrAlaMetValPheLysProHisAsnPheMetLeuGlyValLysTyrPro 515  
QY 992 ACTGCTGCTCATATATGCTGCTCTCAG----- 1021  
Db ThrValLeuAsnValTyrGlyGlyProGluValGlnThrValSerAsnThrPheLysGly 535

QY 1021 ----- 1021  
Db MetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyrCysValIleCysValAsp 555  
QY 1021 ----- 1021  
Db SerArgGlySerArgHisArgGlyValGluPheGluSerTyrIleArgCysArgMetGly 575  
QY 1021 ----- 1021  
Db ThrValGluLeuSerAspGlnValGluValLeuArgIleLeuAlaAspGlnLeuGlyTyr 595  
QY 1021 ----- 1021  
Db IleAspMetAspArgValAlaIleHisGlyTyrPheSerTyrGlyGlyTyrLeuSerLeuMet 615  
QY 1022 -----GTTTGTATTGCTGGGGCCCTCAGTCTCTG 1051  
Db GlyLeuValGlnTyrProGluIlePheLysValSerIleAlaGlyAlaProValThrSer 635  
QY 1052 TGGATCTTCTATGATACAGGATACACGAACTTATATGGTCCACCTGACCGAATGAA 1111  
Db TrpGluTyrTyrAspThrGlyTyrThrGluArgTyrMetAspLeuProAspSerAsnArg 655  
QY 1112 CAGGCTATTACTTACGATCTGTGGCATCGCAAGCAAGAAATTCCTCTGAACCAAT 1171  
Db SerGlyTyrAlaIleGlySerValLeuAsnTyrIleGlnLysPheProAspGluAspAsn 675  
QY 1172 CGTTTACTGCTCTTATCATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1231  
Db ArgLeuLeuIleIleHisGlyLeuIleAspGluAsnValHisPheHisThrSerGln 695  
QY 1232 TTACTGATGTTTTTGTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAG 1291  
Db LeuValSerArgLeuValArgAlaAsnLysProTyrGlnLeuGlnValTyrProAsnGlu 715  
QY 1292 AGACACAGCATAAGATGTTCTGAAATCGGAGAACATATATGAACTCATCTTTTGCACATC 1351  
Db ArgHisSerLeuArgAsnLeuGluAlaSerLysHisTyrGluThrLysLeuSerPhe 735  
QY 1352 CTTCAAGAAACCTT 1366  
Db LeuGlnAsnHisLeu 740  
RESULT 11  
Q9VC20 DROME PRELIMINARY; PRT; 1053 AA.  
AC Q9VC20, OBIH07;  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE G3744-PA, isoform A (G3744-PC, isoform C) (L033755p).  
GN ORFNames:G3744;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidae; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallo R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,  
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beson K.Y., Benos P.V., Berwan B.P., Bhargava D., Bolshakov S.,  
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,



Db 631 ValThrAlaSerLeuLeuLeuSerGln-----AlaAenGlyGlnProAspPro 646  
QY 506 -----AGTGATTTCAAGTGCCTATCAAA 529  
Db 647 GlySerValGlyAlaGlnProSerPheValAspLeuSerAlaLeuGlnProArgIleLeu 666  
QY 530 GAGGAGATACGAATTACCACTGCTGTAATGGAGTCTTGGCGGCATGGATCAATATC 589  
Db 667 AsnLysValAlaLeuThrSerGlyGlyTrpGluValLeuAlaArg-----AsnLeu 683  
QY 590 CAGTTTCATGAAGTCAGAGCGTGGTATATTTGAAGCACCAAGACTCCCTTTAGAG 649  
Db 684 TrpValAspLysAlaLeuValThrSerGlyGlyLeuValArgAspThrProLeuGlu 703  
QY 650 CATCACTGATAGTACGTACGTACGTAACATCTCGAGAGGTGACAAAGCGCTACGTGACCGT 709  
Db 704 LysHisLeuTyrValValSerLeuGluArgProGluHisIleArgLeuLeuThrGluPro 723  
QY 710 GGTCTACTACATTTCTTGC-----TGCACTCAGT 736  
Db 724 GlyTyrSerTyrLeuValGluPheAspAspHisPheAsnAspLeuPheProIlePheSer 743  
QY 737 CAGCACTGTGACTCTTTTATAGTAAGTATAGTAAC---CAGAAAGATCCACACTGTGTG 793  
Db 744 GlnGlnCysLysLeuMetLeuLeuValTyrCysAsnIleGlnArgLeuProSerCysLys 763  
QY 794 TCCCTTTACAAGCTATCAAGTCTCAAGATGACCCCACTGCAAAACAAAGGAATTTGG 853  
Db 764 ValMet-----ArgValaenGlnThrCysSerAenGlyGlyValaen 777  
QY 854 GCCACCATTTGGATTCAGCAGCT-----CCTCTCTCGACTAT 892  
Db 778 GlyIleGlnIleSerLeuValGlyTyrLeuHisGluGlyGlyLysProGluProGlnTyr 797  
QY 893 ACTCTCCGAAATTTCTTTTGAAGTACTACTGATTCATGATTACATTTGATGGATGCTC 952  
Db 798 Cys---ProGlnIlePheSerProGlnLeuProSerGlyAspIleValTyrAlaMetVal 816  
QY 953 TACAGGCTCATGATCTACGCTCGAAGAAATATCTACTGTGCTGTCATATAGT 1012  
Db 817 PheLysProHisAsnPheGluLeuGlyValLysTyrProThrValLeuAenValTyrGly 836  
QY 1013 GGTCTCTCAG----- 1021  
Db 837 GlyProGluValGlnThrValaenAsnThrPheLysGlyLysHisGlnLeuArgMetHis 856  
QY 1021 ----- 1021  
Db 857 MetLeuAlaAlaGlnGlyTyrCysValIleCysIleAspSerArgGlySerArgHisArg 876  
QY 1021 ----- 1021  
Db 877 GlyLysArgPheGluSerHisIleArgGlyArgMetGlyGlnValGluLeuThrAspGln 896  
QY 1021 ----- 1021  
Db 897 ValAspAlaLeuArgSerLeuSerAspGlnLeuGlyTyrIleAspMetAspArgValAla 916  
QY 1021 ----- 1021  
Db 917 IleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProLys 936  
QY 1022 -----GTTGCTATTCTGGCGCCAGTCACTCTGGATCTCTATGATACAGA 1072  
Db 937 IlePheLysValAlaIleAlaGlyAlaProValThrAsnTrpGluTyrTyrAspThrGly 956  
QY 1073 TACACGGAAGTATTATGGTCCCTGACCAAGATGAACAGGCGTATTACTTAGATCT 1132  
Db 957 TyrThrGluArgTyrMetAspMetProGlnaenAsnGluAlaGlyTyrSerAlaGlySer 976  
QY 1133 GTGGCCATGCAAGCAGAAAGTCCCTCTCGAACCAATCGTTTACTGCTCTTACATGTT 1192  
Db 977 ValLeuGluTyrValaenSerPheProGluGluAspLysArgLeuLeuLeuIleHisGly 996

QY 1193 TTCCTGGATGCAGATCTCCATTTTGCATACAGTATATTTACTAGTITTTTACTGAGG 1252  
Db 997 LeuIleAspGluAenValHisPheCysHisThrSerArgLeuIleSerAlaLeuAenLys 1016  
QY 1253 GCTGGAAGCCATATGATTTTACAGATCTTCTCTCAGAGAGACACAGCATAAAGATTCCT 1312  
Db 1017 AlaAenLysProTyrGluValHisLeuPheProGluGluArgHisSerLeuArgAenLeu 1036  
QY 1313 GAATCGGGAGAACATTAATGAATGTCATCTTTCGACTACCTTCAA 1357  
Db 1037 GluSerAenLysAenTyrGluThrLysLeuLeuSerPheLeuGln 1051  
RESULT 12  
ID Q9VC19\_DROME PRELIMINARY; PRT; 1113 AA.  
AC Q9VC19;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE CG3744-PB, isoform B.  
GN Name=CG3744; ORFNames=CG3744;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Wan K.H., Doyle C., Baxter E.G., Heit J.G., Nelson C.R., Miklos G.L.G.,  
RA Abell J.F., Agbayani A., An H.-J., Andrews-Flannkoc C., Baldwin D.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berwan B.P., Bhakdevi P., Brottier P.,  
RA Borokva D., Botchan M.R., Bouck J., Brokstein D., Bolshakov S.,  
RA Churris J.M., Cawley S., Dahike C., Cadieu E., Center A., Chandra I.,  
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwa C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshtafa A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,  
RA Patzold M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson R.D.C., Scheeler F., Shen H.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svrskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195 (2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celnik S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,



Db 917 MetLeuAlaAlaGlnGlyTyrCysValIleCysIleAaspSerArgGlySerArgHisArg 936  
Qy 1021 -----  
Db 937 GlyLysArgPheGluSerHisIleAlaGlyArgMetGlyGlnValGluLeuThrAspGln 956  
Qy 1021 -----  
Db 957 ValAspAlaLeuArgSerLeuSerAspGlnLeuGlyTyrIleAaspMetAspArgValAla 976  
Qy 1021 -----  
Db 977 IleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProLys 996  
Qy 1022 -----GTTCTATTCTGGGCGCCAGCTCCTCTGTGATCTTCTATGATACAGGA 1072  
Db 997 IlePheLysValAlaIleAlaGlyAlaProValThrAsnTrpGluTyrTyrAspThrGly 1016  
Qy 1073 TACACGAGAGCTTATATGGTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCT 1132  
Db 1017 TyrThrGluArgTyrMetAspMetProGlnAenAenGluAlaGlyTyrSerAlaGlySer 1036  
Qy 1133 GTGCCCATGCAACGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGT 1192  
Db 1037 ValLeuGluTyrValAenSerPheProGluGluAspLysArgLeuLeuIleHisGly 1056  
Qy 1193 TTCTGTGATGAGATGCTCATTTTGCACATACCACTATATTTACTGAGTTTTTTAGTCAGG 1252  
Db 1057 LeuIleAspGluAenValHisPheCysHisThrSerArgLeuIleSerAlaLeuAenLys 1076  
Qy 1253 GCTGGAAAGCCATATGATTACAGATCTATCTCTCAGGAGACACACAGATAGAGTTTCTT 1312  
Db 1077 AlaAenLysProTyrGluValHisLeuPheProGluGluArgHisSerLeuArgAenLeu 1096  
Qy 1313 GAATCGGAGAACATTATGATGATCTTTTGTGACCTACCTTCAA 1357  
Db 1097 GluSerAsnLysAenTyrGluThrLysLeuLeuSerPheLeuGln 1111  
RESULT 13  
QSTXJ2 ANOGA  
ID QSTXJ2 ANOGA PRELIMINARY; PRT; 557 AA.  
AC QSTXJ2;  
DT 01-FEB-2005 (T-EMBLrel. 29, Created)  
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)  
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)  
DE ENSANGP00000029249 (Fragment).  
GN ORFNames=ENSANGG0000012958;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;  
OC Anophelinae; Anopheles.  
OX NCBI\_TaxID=180454;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=PEST;  
RG The Anopheles gambiae Sequence Committee;  
RT "Anopheles gambiae re-annotation";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN (2)  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=PEST;  
RG The Anopheles gambiae Sequence Committee;  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
CC -! CAUTION, The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AA001008287; EAL42106.1; -; Genomic DNA.  
DR GO; GO:0003824; F: catalytic activity; IEA.  
DR GO; GO:0008233; F: serine-type peptidase activity; IEA.  
DR GO; GO:0008506; F: proteolysis and peptidolysis; IEA.  
DR InterPro; IPR0011375; Peptidase\_S9.  
DR InterPro; IPR002469; Peptidase\_S9B.

InterPro; IPR000379; Ser\_estrs.  
DR Pfam; PF00930; DPPIV\_N; I.  
DR Pfam; PF00326; Peptidase\_S9; 1.  
FT NON\_TER 557 1  
FT NON\_TER 557 1  
SQ SEQUENCE 557 AA; 62685 MW; 08AAA0944ACD0808 CRC64;  
Alignment Scores:  
Pred. No.: 1.55e-51 Length: 557  
Score: 718.50 Matches: 185  
Percent Similarity: 43.9% Conservative: 56  
Best Local Similarity: 33.7% Mismatches: 159  
Query Match: 24.3% Indels: 149  
DB: 2 Gaps: 13  
US-10-825-632-6 (1-1669) x QSTXJ2\_ANOGA (1-557)  
Qy 5 GGTACACCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATTAATGATTGATGCTGAAGGA 64  
Db 36 GlyThrProAenAlaLysSerLysLeuValGlnPheArgLeuSerGluAenLeu 55  
Qy 65 AGGATCATAGATGTCATAGATAAGGAACCTAATCAACCTTTTGAGATCTCTATTGTAAGGA 124  
Db 56 ArgIleThrAspValCysIleLysGluLeuGlnCysProLeuThrPheAlaPheProTyr 75  
Qy 125 GTTGAATATATGCCAGAGCTGGATGGACTCTCTGAGGAAATATATGCTTGCTGCTCATCTA 184  
Db 76 LeuGluTyrIleValArgValGlyTyrThrProAspSerArgTyrValTrpAlaGlnLeu 95  
Qy 185 CTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACCTGAATATTATATCCCA 244  
Db 96 LeuAspArgProGlnArgLeuGluLeuValLeuLeuProValAspAenPheCysSer 115  
Qy 245 GTAGAGATGATGTTATGGAAGGCAGAGACTCATTGAGTCAGTGCCTGATCTTGTGACG 304  
Db 116 ThrAla-----AlaArgArgArgProThrAlaGlyArgProProGlyAlaMet 133  
Qy 305 CCACCTAATATCTTCAAGAAACACACATCTGGATAAATATCCATGACATCTTTCAT 364  
Db 134 GlyProGlyIleTyrThrGluThrSerSerTyrValAenValHisAspValLeuGln 153  
Qy 365 GTTTTCCCAAGTCACGAGAGAAATTCGTTTATTTTGGCTCTGAAATGCAAAACA 424  
Db 154 ---PheValGluLeuSerGluGlnGluVal-----Thr 163  
Qy 425 GGTTCCTGCATTTATACAAATTAATCATCTATTTTAAAGGAAGCAAAATAT-----AAA 478  
Db 164 ThrAenSerHisProThrAlaHisProLysGlnArgGluGlyAlaHisGlyThrGlu 183  
Qy 479 CGATCCAGTGGTGGCTGCTGCTCCAGTCAAGTCAATTTCAAGTGTCT----- 523  
Db 184 valSerSerGlyGlyAlaGlyAla---ThrAspHisSerLeuProSerMetAlaCysIle 202  
Qy 524 -----ATCAAGAGAGAGATAGCAATTAACCATGCTGCTGAATGGAA 562  
Db 203 GlySerThrLeuValAlaArgIleValGlnLysValThrLeuThrGlyGlyAspTrpGlu 222  
Qy 563 GTTCTTGCCCGCATGGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTT 622  
Db 223 ValLeuGlyArg-----AsnValTrpTyrAspArgValArgGlnLeuValTyrPhe 239  
Qy 623 GAAGGCACCAAGACTCCCTTTAGACATCACCTGTACGTACGTACGTACGTAAATCCT 682  
Db 240 MetGlyLeuArgGluThrProLeuGluLysHisLeuTyrValValSerLeuAlaGlnPro 259  
Qy 683 GGAGAGGTGACCAAGGCTGACTGACCTGGCTACTCATTCTTCTGCTGATCAGTCAGCAC 742  
Db 260 AenGlnLeuArgLeuLeuThrMetProGlyTyrSerPheThrVal-----GluPhe 276  
Qy 743 TGTGACTTCTTT---ATAAGTAAGTATAGTAACACAGAAATCCACACTGTGTGCTCCTT 799  
Db 277 AenAspLeuValArgIleAlaHisAspSerAsnThrAlaAenGlyAenGlyCysSer--- 295

```
QY 800 TACAAGCTATCAAGTCTCGAAGATGACCCAACTTCGAAACAAAGGAATTTTGGCCACC 859
Db 296 -----HisGlyProThrProProProThr-----ProIleAspAlaLeu 308
QY 860 ATTTTGATTCACAGAGTCTCTCTCTGACTATCTCT----- 898
Db 309 ArgLeuCysSerValGlyTyrLeuThrGluGlyGlyProSerGluAsnThrGlnTyrAsn 328
QY 899 CCAGAAATTTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATCGCTACAG 958
Db 329 ProSerIleHisSerProGlnIleSerSerGlyAspValLeuTyrAlaMetValPheLys 348
QY 959 CTTTCATGATCTACAGCTGCGAAGAAATATCTCTGCTGCTTCATATATGGTGGCCT 1018
Db 349 ProHisAsnPheMetLeuGlyValLysTyrProThrValLeuAsnValTyrGlyPro 368
QY 1019 CAG----- 1021
Db 369 GluValGlnThrValSerAsnThrPheLysGlyMetArgGlnLeuArgMetHisMetLeu 388
QY 1021 ----- 1021
Db 389 AlaSerGlnGlyTyrCysValIleCysValAspSerArgGlySerArgHisArgGlyVal 408
QY 1021 ----- 1021
Db 409 GluPheGluSerTyrIleArgArgMetGlyThrValGluLeuSerAspGlnValGlu 428
QY 1021 ----- 1021
Db 429 ValLeuArgIleLeuAlaAspGlnLeuGlyTyrIleAspMetAspArgValAlaIleHis 448
QY 1021 ----- 1021
Db 449 GlyTrpSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProGluIlePhe 468
QY 1022 ---GTTCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATFACAGGATACACG 1078
Db 469 LysValSerIleAlaGlyAlaProValThrSerTrpGluTyrTyrAspThrGlyTyrThr 488
QY 1079 GAACGTTATATGGTCAACCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCC 1138
Db 489 GluArgTyrMetAspLeuProAspSerAsnArgSerGlyTyrAlaAlaGlySerValLeu 508
QY 1139 ATGCACACGAAAGTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTGTG 1198
Db 509 AsnTyrIleGlnLysPheProAspGluAspAsnArgLeuLeuIleHisGlyLeuIle 528
QY 1199 GATCAGAAATCTCCATTTTGCACATACAGTATATTACTAGTCTTTTGTAGCGGCTGGA 1258
Db 529 AspGluAsnValHisPheHisIleThrSerGlnLeuValSerArgLeuValArgAlaAsn 548
QY 1259 AAGCCATATGATTTACAGATCTATCTCT 1285
Db 549 LysProTyrGlnLeuGlnValTyrPro 557
```

## RESULT 14

```
Q54U01.DICDI
ID Q54U01.DICDI PRELIMINARY; PRT; 803 AA.
AC Q54U01;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD80205566;
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_taxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
```

```
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsey R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Lousaged H., Mungall K., Oliver K., Price C., Quail M.A., Sanders M.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noeig A.A., Barrell B., Kuspa A.,
RT "The genome of the social amoeba Dictyostelium discoideum."
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC -!- ENBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF101000074; EAL66689.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 803 AA; 93187 MW; 33846B370C039FAE CRC64;
```

## Alignment Scores:

```
Pred. No.: 1.27e-42 Length: 803
Score: 613.50 Matches: 166
Percent Similarity: 43.0% Conservative: 79
Best Local Similarity: 29.1% Mismatches: 160
Query Match: 20.7% Indels: 165
DB: 2 Gaps: 18
```

US-10-825-632-6 (1-1669) x Q54U01.DICDI (1-803)

```
QY 20 AAATCTCTTTTAAAGATCTCAGAAATAATGATTGATGCTGAAGGAGGATCATAGATGTC 79
Db 278 ArgThrThrPheGlnAspSerLys-----IleGluIle 288
QY 80 ATAGATAAGCACTAATTCACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATGCC 139
Db 289 ValLysSerGluLeuPheAsp---LeuLysThrGlnPheProTrpAlaGluTyrIleThr 307
QY 140 AGACCTCGATGAGTCTCTCTGAGGAAATATGCTTGCTCCATCTCTACTAGATCGTCCAG 199
Db 308 ArgAlaGlyTrpThrProAsnGlyHisSerIleTyrLeuGlnLeuLeuAspArgLysGln 327
QY 200 ACTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATTT 259
Db 328 GlnHisLeuAlaLeuValMetValProLeuHisValPhe-----AlaGluAspTyrSer 345
QY 260 ATGAAAGGAGGAGACACTCATTTAGTGTGCTGATTTCTGTGACGCCACTAATTTCTAT 319
Db 346 SerSerSerSerSerValLysSerIleProLys-----LeuProValLeuIle--- 362
QY 320 GAAGAAACACAGACATCTGGATAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGT 379
Db 363 GluGluThrThrSerValTrpIleAsnIleGluPheSerPheGln---PheLeuLysSer 381
QY 380 CACGAGAGGAAATTTGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTCGTCATTTA 439
Db 382 IleGluAsnGlnLeu-----IleTrpSerAsnGlu---GlnSerGlyTyrArgHisLeu 398
QY 440 TACAAAATATCATCTATTTTAAAGAAAGCAATATAAACGATCCAGTGTGGTGGCTGCT 499
Db 399 Tyr-----LeuIleLysTrpAspLysAsnPheThrAsnIleGlnSerThrPro 414
QY 500 GCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGGAGAGATAGCAATATACAGTGGTGAATGG 559
Db 415 IleThrLeuSer---ThrCysAsnAspAsnAspAsn-----AspAsnAsnAsnTrp 430
QY 560 GAAGTTCTTGGCCGGCGATGGATCTATATATCCAAAGTTGATGAAGTACAGAGGCTGTATAT 619
```



```
Db 431 MetVal-----SerSerAspAspIleHisIleAspGluLysArgLysLeuValTyr 447
QY 620 TTTGAGGACCAAGACTCCCTTTAGAGCATCACCCTGACGTAGTACGTAAAT 679
Db 448 PheThrGlyThrLysAspThrCysLeuGluGlnHisLeuTyrValThrArgPheAspLys 467
QY 680 CCT---GGAGAGTGACAGGCTGACTGACCGTGGCTACTACAT---TCTTGCTGCATC 733
Db 468 ProAsnSerGluIleLysArgLeuSerHisAlaAsnPheSerHisArgSerIleSerIle 487
QY 734 AGTCAGCAGTGTGACTTCTTTATAGTAAGTATAGTAAACCAAGAAGAT----- 781
Db 488 SerSerAsnPheLysPheIleThrThrThrThrThrThrThrThrThrThrThr 507
QY 782 -----CCACAC 787
Db 508 GluValPheAspLeuIleTyrAsnAspAspAsnAsnAspAspAspAspIleTyrProIle 527
QY 788 TGTGTGTCCCTTTACAGCTATCAAGTCTCCTGAAGATGACCCCAACTTGCMAAACAAAGAA 847
Db 528 ValLysSerSerPhePheIleAsnAspAspAspAspAspAspAspAspAspLysLys 547
QY 848 TTTTGGGCCACCACTTTTGGATTTCAGCAGTCTCTTCTGCTACTACTCTCCAGAAATT 907
Db 548 IleAsnIleAsnIle-----ProLysIle 555
QY 908 TTCTCTTTCAAGTACTACTGATTTACATTGTATGGATGCTCTACAGCCTCATGAT 967
Db 556 PheAsnPheLysAsnSerLysGlyValThrIleTyrGlyGlnTyrThrLeuProSerAsp 575
QY 968 CTACAGCTGGAAGAAATCTACTGTCTGTTCATATATATGTTGCTCT----- 1018
Db 576 TyrSerLysAspLysIleProThrValValTyrValTyrValTyrValTyrValGln 595
QY 1018 ----- 1018
Db 596 IleValArgAsnGlnTyrAsnTyrIleLysGlnHisTyrThrAsnPheGlyPheIleGln 615
QY 1018 ----- 1018
Db 616 ValMetIleAspAsnValGlySerAlaAsnArgGlyLeuGluPheGluSerHisIleArg 635
QY 1018 ----- 1018
Db 636 GluLysMetGlyGlnValGluIleGlyAspGlnValGluGlyIleAsnTyrLeuValGly 655
QY 1018 ----- 1018
Db 656 AsnAspIleValSerIleAspValAsnArgIleAlaIleSerGlyTyrSerTyrGlyGly 675
QY 1019 -----CAGGTTGCTATTGCTGGG 1036
Db 676 TyrAsnSerLeuMetAlaIleSerGlnArgProAspValPheLysIleAlaValCysGly 695
QY 1037 GCCCCAGTCACTCTGTGGATCTTATGATACAGGATACAGGAACGTTATATGGTTCAC 1096
Db 696 AlaProValSerAspTrpArgLeuTyrAsnThrGlyTyrThrGluArgTyrMetAsnVal 715
QY 1097 CTGACCAAGATGAACAGGCTATTACTAGGATCTGGCCATGTCGCAAGCAGAAAGTTC 1156
Db 716 ProGlnAspAsnIleAspGlyTyrLysLeuGlyAspThrHisTyrSer-----Phe 733
QY 1157 CCTCTCAACCAATCGTTACTGCTTATCATGTTCTCGGATGATGATGCTCAATTT 1216
Db 734 ProThrGluLysArgLeuLeuIleHisGlyLeuGlnAspGluAsnValHisPhe 753
QY 1217 GCATACCAAGTATATTACTAGTATTTTATGAGGCTGGAAGCCATATGATTTACAG 1276
Db 754 SerAsnThrIleTyrIleAspHisLeuThrLysThrGlnLysProTyrIleLeuLys 773
QY 1277 ATCTATCTCAGAGACACAGCATATAGATTTCTGATCGGGAGAACATTATGAACTG 1336
Db 774 ThrLeuProAsnGluArgHisGlyValArgAsnThrAspAsnArgIleTyrIleGlyLeu 793
```

```
QY 1337 CATCTTTTGCACTACCTTCAAGAAAAACCTT 1366
Db 794 PheValIleAsnHisLeuLysAsnLeu 803
RESULT 15
Q965K3_CABEL
ID Q965K3_CABEL PRELIMINARY; PRT; 927 AA.
AC Q965K3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dipeptidyl peptidase four (IV) family protein 3, isoform b.
GN Name=dpf-3; ORFNames=K02F2.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=95069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology.";
RL Science 282:2012-2018 (1998).
DR EMBL AF043699; AAR84627.1; -; Genomic_DNA.
DR MEROPS; S09.016; -;
DR Ensembl; K02F2.1; Caenorhabditis elegans.
DR WormBase; WSGene0001056; dpf-3.
DR WormPept; K02F2.1b; CE28837.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:diacylglycerol 3-phosphate 4-epimerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S98.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR Complete proteome.
SQ SEQUENCE 927 AA; 105366 MW; 6907C6AEF829D0F CRC64;
```

## Alignment Scores:

Pred. No.:	8,22e-30	Length:	927
Score:	462.00	Matches:	154
Percent Similarity:	38.7%	Conservative:	79
Best Local Similarity:	25.6%	Mismatches:	177
Query Match:	15.6%	Indels:	192
DB:	2	Gaps:	23

US-10-825-632-6 (1-1669) x Q965K3\_CABEL (1-927)

```
QY 5 GGTACAGCAATCCTAAATCCTTTTAAAGATGTCAGAAATATGATGCTGAAGGA 64
Db 351 GlyThrLysAsnAlaTyrSerThrLeuArgMet-----ValIleLeuGluAsnGly 367
QY 65 AGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATCTCTATTGAAGGA 124
Db 368 LysAlaTyrAspValProLeuLysAsp-----GluValIleTyrLysHis 382
QY 125 -----GTTGATATATATTCCTCCAGACTCGATGAGTCTCTGAGGAAATATGCTTGG 175
Db 383 CysProPheTyrGlnTyrIleThrArgAlaGlyPhePheSerAspGlyThrThrValTrp 402
QY 176 TCCATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATC----- 223
Db 403 ValGlnValMetSerArgAspGlnAlaGlnCysSerLeuLeuLeuIleProTyrThrAsp 422
QY 224 -----TCACCT 229
Db 423 PheLeuLeuProGluLeuGlyGlySerIleLysGluAspAsnLeuLeuLeuSerThr 442
```



```
QY 230 GAATTATTTATCCAGTAGAATGATGTTATGAAAGGACAGACACTCATTTGAGTCAGTG 289
Db 443 ApeLeuAsnMetGlyValTrpAspAsp---LysSerHisGluGluThrMetGluLysPro 461
QY 290 CTTGATCTGTGAGC-----CCACTAATTTCTATGAGAAACACACAGACATCTCGATA 343
Db 462 ProArgGlyLysLeuArgGlyThrValGlnIleHisLysAlaArgAsnAspTyrTrpIle 481
QY 344 AATATCATGACATCTTTTCATGTTTCCCAAGTCCAGAAAGGAAA-----ATTGAG 397
Db 482 AsnThrHisAsnAlaIleTyrProLeuLysIleThrAspGluGluHisProMetTyrGlu 501
QY 398 TTTATTTTTCCTCTGAATGCAAAACAGGTTTCCTCATTTATACAAAATATCATCTATT 457
Db 502 PheIleTyrCysLeuGluLysProAsnGly--SerCysLeuAlaLeuIleSerAlaGlu 520
QY 458 TTAAGGAAGCAATATATAACGATCCAGTCGTGGTGGCTGCTCCCAAGTGATTTCAAG 517
Db 521 LeuAspGlnAsnGlyTyr----- 526
QY 518 TGTCCCTATCAAGAGGAGATACGAATTACCACTGCTGTAATGGGAAGTTCTTGGCCGCAT 577
Db 527 CysArgHisThrGluGluLysLeuLeuMetAlaGluAsnPheSerIle-----AsnLys 544
QY 578 GATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAC 637
Db 545 SerMetGlyIleValValAspGluValAspGluLeuValTyrTyrValAlaAsnGluSer 564
QY 638 TCCCTTTAGACATCACCTGTACGTAGTACGTAGTAAATCTCTGGAGAGGTGACAAAG 697
Db 565 HisProThrGluTrpAsnIle---CysValSerHisTyrArgThrGlyGlnHisAlaGln 583
QY 698 CTGACCTGACCGTGGCTACTCACATCTTGTGTCATCAGTCAGTCAGTCTGCTTTTATA 757
Db 584 LeuThrGluSerGlyIle-----Cys-----PheLys 592
QY 758 AGTAAGTATAGTAACACAGAG-----AATCCACACTGTGTGTCCTTTTACAAAG 805
Db 593 SerGluArgAlaAsnGlyLysLeuAlaLeuAspLeuAspHisGlyPheAlaCysTyrMet 612
QY 806 CTATCA-----AGTCCTGAAGATGACCCCACTTCCAAACAAAGGAATTT---TGG--- 853
Db 613 ThrSerValGlySerProAlaGlu-----CysArgPheTyrSerPheArgTrpLys 629
QY 854 -----GCCACCATTTTGGATTTCAGCAGGTCTCT 880
Db 630 GluAsnGluValLeuProSerThrValTyrAlaAlaAsnIleThrValSerGlyHisPro 649
QY 881 ---CTTCCTGAC-----TATACTCTCCAGAAATTTTCTCTTTTGAAGT---ACTACT 928
Db 650 GlyGlnProAspLeuHisPheAspSerProGluMetIleGluPheGlnSerLysLysThr 669
QY 929 GGATTTACATTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGGAAAGAAATAT 988
Db 670 GlyLeuMetHisTyrAlaMetIleLeuArgProSerAsnPheAspProTyrLysLysTyr 689
QY 989 CCTACTGCTGCTCATATATGTTGCTCT----- 1018
Db 690 ProValPheHisTyrValTyrGlyGlyProGlyIleGlnIleValHisAsnAspPheSer 709
QY 1018 ----- 1018
Db 710 TrpIleGlnTyrIleArgPheCysArgLeuGlyTyrValValValPheIleAspAsnArg 729
QY 1018 ----- 1018
Db 730 GlySerAlaHisArgGlyIleGluPheGluArgHisIleHisLysLysMetGlyThrVal 749
QY 1018 ----- 1018
Db 750 GluValGluAspGlnValGluGlyLeuGlnMetLeuAlaGluArgThrGlyGlyPheMet 769
QY 1018 ----- 1018

Db 770 AspMetSerArgValValValHisGlyTrpSerTyrGlyGlyTyrMetAlaLeuGlnMet 789
QY 1019 -----CAGGTTGCTATTGCTGGGGCCCGCCAGTCACCTCTCTGG 1054
Db 790 IleAlaLysHisProAsnIleTyrArgAlaAlaIleAlaGlyGlyAlaValSerAspTrp 809
QY 1055 ATCTTCTATGATACAGGATACACGAAAGCTTATATGGTCCACCTGACCAAGAAATGAACAG 1114
Db 810 ArgLeuTyrAspThrAlaTyrThrGluArgTyrMetGlyTyrPro---LeuGluGluHis 828
QY 1115 GGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGATTCCCTCTGAAACCAAAATCGT 1174
Db 829 ValTyrGlyAlaSerSerIleThrGlyLeuValGluLysLeuProAspGluProAsnArg 848
QY 1175 TTACTGCTCTTACATCGTTTCTCTGGATGAGAAATGTCCATTTTGCACATACCAAGTATATTA 1234
Db 849 LeuMetLeuValHisGlyLeuMetAspGluAsnValHisPheAlaHisLeuThrHisLeu 868
QY 1235 CTGAGTTTTTTAGTAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGA 1294
Db 869 ValAspGluCysIleLysLysGlyLysTrpHisGluLeuValIlePheProAsnGluArg 888
QY 1295 CACAGCATAGAGTTCTCTGAATCGGAGAGAACATTATGAACTGTCATCTTTTGCACCTACCTT 1354
Db 889 HisGlyValArgAsnAsnAspAlaSerIleTyrLeuAspAlaArgMetMetTyrPheAla 908
QY 1355 CAAGAA 1360
Db 909 GlnGln 910

Search completed: May 2, 2006, 01:33:18
Job time : 424.39 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 01:34:19 ; Search time 10.2279 Seconds  
(without alignments)  
2698.234 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 2957  
Sequence: 1 aaacgtagcagcaaatccta.....aaaaaaaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abss/ABSSWEB\_spool/US10825632/runat\_01052006\_105950\_3311/app\_query.fasta\_1  
-DB=Issued Patents AA -QFM=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abss02p -USER=US10825632 @CGN 1 1 131 @runat\_01052006\_105950\_3311  
-NCFU=6 -ICFU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5 COMB pep:\*  
2: /cgn2\_6/prodata/1/iaa/6 COMB pep:\*  
3: /cgn2\_6/prodata/1/iaa/H COMB pep:\*  
4: /cgn2\_6/prodata/1/iaa/PCTUS COMB pep:\*  
5: /cgn2\_6/prodata/1/iaa/RE COMB pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2482	83.9	465	2	US-10-070-464-5
2	2422	81.9	882	2	US-09-976-674-1
3	2422	81.9	882	2	US-10-070-464-1
4	1957.5	66.2	690	2	US-09-976-674-7
5	1808	61.1	658	2	US-09-976-674-19
6	1808	61.1	661	2	US-09-976-674-11
7	1540	52.1	613	2	US-09-976-674-21
8	1494.5	50.5	863	2	US-09-976-674-3
9	1494.5	50.5	882	2	US-09-976-674-23
10	1494.5	50.5	892	2	US-09-976-674-27
11	1445	48.9	879	2	US-09-976-674-33
12	1445	48.9	879	2	US-09-976-674-35

13	1377.5	46.6	310	2	US-09-794-236-4	Sequence 4, Appli
14	1377.5	46.6	310	2	US-10-070-464-3	Sequence 3, Appli
15	1094.5	37.0	832	2	US-09-976-674-29	Sequence 29, Appl
16	1094.5	37.0	832	2	US-09-976-674-31	Sequence 31, Appl
17	1045	35.3	819	2	US-09-976-674-37	Sequence 37, Appl
18	1045	35.3	819	2	US-09-976-674-39	Sequence 39, Appl
19	860.5	29.1	360	2	US-10-070-464-7	Sequence 7, Appli
20	471.5	15.9	518	2	US-09-976-674-25	Sequence 25, Appl
21	309.5	10.5	771	2	US-09-462-584-2	Sequence 2, Appli
22	309.5	10.5	771	2	US-09-079-592-2	Sequence 2, Appli
23	305	10.3	676	2	US-09-518-550-42	Sequence 42, Appl
24	305	10.3	723	2	US-09-518-550-29	Sequence 29, Appl
25	296.5	10.0	702	2	US-09-518-550-28	Sequence 28, Appl
26	294	9.7	102	2	US-09-270-767-44517	Sequence 44517, A
27	288	9.7	732	2	US-09-518-550-30	Sequence 30, Appl
28	281.5	9.5	981	2	US-09-902-540-16812	Sequence 16812, A
29	277	9.4	818	2	US-09-462-845-3	Sequence 3, Appli
30	277	9.4	818	2	US-10-402-312-3	Sequence 3, Appli
31	277	9.4	818	2	US-10-401-437-3	Sequence 3, Appli
32	277	9.4	818	2	US-10-402-067-3	Sequence 3, Appli
33	277	9.4	818	2	US-10-401-436-3	Sequence 3, Appli
34	267.5	9.0	931	2	US-09-079-592-11	Sequence 11, Appl
35	263	8.9	54	2	US-09-621-976-5647	Sequence 5647, Ap
36	249.5	8.4	706	2	US-09-976-674-41	Sequence 41, Appl
37	249.5	8.4	743	2	US-10-363-937-4	Sequence 4, Appli
38	249.5	8.4	796	2	US-09-976-674-5	Sequence 5, Appli
39	244	8.3	737	2	US-09-902-540-11421	Sequence 11421, A
40	238	8.0	766	2	US-10-002-593-6	Sequence 6, Appli
41	238	8.0	766	2	US-09-949-016-6146	Sequence 6146, Ap
42	238	8.0	766	2	US-10-423-714-6	Sequence 6, Appli
43	234	7.9	755	4	PCT-US93-07923-3	Sequence 3, Appli
44	234	7.9	759	4	PCT-US93-07923-2	Sequence 2, Appli
45	234	7.9	766	1	US-08-230-491A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-10-070-464-5  
; Sequence 5, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 465  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-070-464-5

Alignment Scores:  
Pred. No.: 3.76e-283 Length: 465  
Score: 2482.00 Matches: 465  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 83.9% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-6 (1-1669) x US-10-070-464-5 (1-465)

QY 2 ACAGGTACAGCAAAATCTTAAGTCACCTTTTAAGATGTCAGAAATATGATTGATCTGAA 61

Db 1 ThrGlyThrAlaAenProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 20  
Qy 62 GGAAGGATCATAGATGTCTAGATAGGAACCTAATTCACCTTTTGGATTCCTATTGAA 121  
Db 21 GlyArgIleIleAspValIleAspLysGluIleIleGlnProPheGluIleLeuPheGlu 40  
Qy 122 GGAGTTGAATATATCCGAGACTGGATGACTCTCTGAGGAAAATATGCTTGCTCATC 181  
Db 41 GlyValGluTyrIleAlaArgAlaGlyIrpPmrProGluGlyLysTyrAlaIrpSerIle 60  
Qy 182 CTACTAGATCCCTCCAGACTCGCTACAGATAGTGTGATCTCACTCGAATATTATTC 241  
Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 80  
Qy 242 CCAGTAGACATGATTTATGGAAGGCGAGAGACTCATTTAGTCAGTCCCTGATTCGTG 301  
Db 81 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100  
Qy 302 ACGCCACTAATTATCATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361  
Db 101 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAenIleHisAspIlePhe 120  
Qy 362 CATGTTTTCCTCCCAAGTCACGAAGAGAAATTCAGTTTATTTTGGCTCTGAAATGCAA 421  
Db 121 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 140  
Qy 422 ACAGGTTTCCTCATTTATACAAAATTACATCTATTTTAAAGGAAGCAAAATATAACGA 481  
Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 160  
Qy 482 TCCAGTGGTGGCTGCTGCTCCCAAGTCATTTCAAGTCTCTCAAGAGGAGATAGCA 541  
Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 180  
Qy 542 ATTACAGTGGTGAATGGAGTCTCTGGCCGATGATCATTAATATCAAGTTGATGAA 601  
Db 181 IleThrSerGlyValTrpGluValLeuGlyArgHisGlySerAenIleGlnValAspGlu 200  
Qy 602 GTCAGAGGCTGGTATTTTGAAGGACCAAGACTCCCTTTAGAGCATCCCTGTATC 661  
Db 201 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleTyr 220  
Qy 662 GTAGTCAGTTAGTAAATCCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAT 721  
Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240  
Qy 722 TCTGTCTGCATCAGTCAGACCTGTGACTCTCTTATAAGTAAATAGTAACAGAGAGAT 781  
Db 241 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAen 260  
Qy 782 CCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTCTGAAGTACACCACTTGCAAAACA 841  
Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 280  
Qy 842 AAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTCTCTCACTATATCTCTCCA 901  
Db 281 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300  
Qy 902 GAAATTTTCTCTTTTGAAGTACTACTGGATTTTACATTGTATGGGATGCTCTACAGCT 961  
Db 301 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320  
Qy 962 CATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTCATATATGCTGGTCTCAG 1021  
Db 321 HisAspLeuGlnProGlyLysPheTyrProThrValLeuPheIleTyrGlyGlyProGln 340  
Qy 1022 GTTGCTATTGCTGGGCCCGGACCTGCTGTGGATCTTCTATGATACAGATACAGGAA 1081  
Db 341 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 360  
Qy 1082 GCTTATATGGTACCCCTGACAGAAATGAACAGGGCTATTAATTAGGATCTCTGGCCATG 1141

Db 361 ArgTyrMetGlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMet 380  
Qy 1142 CAACGAGAAAGTTCCCTCTCTGAAACCAATTCGTTTACTGCTCTTACATGGTTTCTCGAT 1201  
Db 381 GlnAlaGluLysPheProSerGluProAenArgLeuLeuLeuHisGlyPheLeuAsp 400  
Qy 1202 GAGAAATCCATTTTGCACATACCACTATATATTACTGAGTTTTTTAGTGAGGCTGGAAAG 1261  
Db 401 GluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 420  
Qy 1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACACAGATAAGAGTTCTCTGAATCGGA 1321  
Db 421 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 440  
Qy 1322 GAACATTATGAACCTGATCTTTTGCATACCTTCAAGAAACCTTGGATCACGTATTGCT 1381  
Db 441 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAla 460  
Qy 1382 GCTCTAAAGTGATA 1396  
Db 461 AlaLeuLysValIle 465  
RESULT 2  
US-09-976-674-1  
; Sequence 1, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akisanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junich, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-1  
Alignment Scores:  
Pred. No.: 6,95e-276 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: 2 Gaps: 1  
US-10-825-632-6 (1-1669) x US-09-976-674-1 (1-882)  
Qy 2 ACAGGTACAGCAAAATCCCTAAAGTCACCTTTAAGATGTCCAGAAATATGATGATGCTGAA 61  
Db 318 ThrGlyThrAlaAenProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337  
Qy 62 GGAAGATCATAGATGTCTAGATAGGAACCTAATTCACCTTTTGGATTCCTATTGAA 121  
Db 338 GlyArgIleIleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGlu 357  
Qy 122 GGAGTTGAATATATTTCCAGAGCTGGATGAGTCTCTCAGGAGAAAATATGCTTGCTCATC 181  
Db 358 GlyValGluTyrIleAlaArgAlaGlyIrpPmrProGluGlyLysTyrAlaIrpSerIle 377  
Qy 182 CTACTAGATCCCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTATTC 241  
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
Qy 242 CCAGTAGAAGATGATGTTTATGGAAGGCGAGAGACTCATTCAGTCAGTCTGCTGTG 301

Db 398 PtoValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 AGCCACTAATTATCTATGAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361  
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleThrAspIleHisAspIlePhe 437  
QY 362 CATGTTTTCCTCCCAAGTCACGAGAGAAATTCAGTTTATTTTGCCTCTGAATCAAA 421  
Db 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGGTTTCCTGTCATTATACAAAATACATCTATTTAAAGGAAAGCAAAATATAACGA 481  
Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
QY 482 TCCAGTGTGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAAAGAGAGATAGCA 541  
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497  
QY 542 ATTACAGTGTGAATGGAGAGTCTTTCGGCCGGCATGGATCTAATATCCAGTTGATGAA 601  
Db 498 IleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTCAGAGCTGGTATATTTCAAGGCACCAAGCTCCCTTTAGAGCATCACCTGTAC 661  
Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537  
QY 662 GTAGTCAGTGTAGTAACTCTCGAGAGGTGACAGGCTGCTGACCGTGGCTACTCACAT 721  
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
QY 722 TCTGTGTCATCAGTCAGTGTGACTTCTTTTAAAGTAAAGTATAGTAAACCAAGAAGAT 781  
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
QY 782 CCACTACTGTGTCCCTTTTCAAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAG 841  
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
QY 842 AAGGAATTTGGCCACCAATTTTGATTCAGAGTCTCTCTCTCTGATATCTCTCTCA 901  
Db 598 LysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
QY 902 GAATTTTCTTTTGAAGTACTACTCGATTTACATTTGATGGATGGATGGATGGATGGAT 961  
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
QY 962 CATGATCTACAGCTCGAAGAAATATCTCTGCTGTCTGCTGCTGCTGCTGCTGCTGCT 1021  
Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657  
QY 1021 ----- 1021  
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677  
QY 1021 ----- 1021  
Db 678 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
QY 1021 ----- 1021  
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
QY 1021 ----- 1021  
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021  
Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTCTGTATGATACAGGATACACGAA 1081  
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTyrPheIlePheTyrAspThrGlyTyrThrGlu 777

QY 1082 CGTTATATGGTCACCTGACCAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGCCATG 1141  
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797  
QY 1142 CAAGCAGAAAAGTTCCCTCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGAT 1201  
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817  
QY 1202 GAGAAATGTCATTTCACATACCATACCATATATATCTGAGTCTTTTCTAGTGGCTGGAAG 1261  
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
QY 1262 CCATATGATTACAGATCTATCTCTCAGGAGACACACATAAGAGTTTCTGGAATCGGA 1321  
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857  
QY 1322 GAACATTATGAATGTCATCTTTTGCATCTCTTCAAGAAAAACCTTGGATCACGTTATGCT 1381  
Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877  
QY 1382 GCTCTAAAAAGTGATA 1396  
Db 878 AlaLeuLysValIle 882  
RESULT 3  
US-10-070-464-1  
; Sequence 1, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-070-464-1  
Alignment Scores:  
Pred. No.: 6,95e-276 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: 2 Gaps: 1  
US-10-825-632-6 (1-1669) x US-10-070-464-1 (1-882)  
QY 2 ACAGGTACACAAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATATGATGCTGAA 61  
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337  
QY 62 GGAAGGATCATAGATGTTCATAGATAAGGAACCTTAATTCACCTTTTTCAGATTCTATTGAA 121  
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357  
QY 122 GGAGTTCAATATATTTGCCAGAGCTGGATGGATCTCTCTGAGGGAAAAATATCTGTGGTCCATC 181  
Db 358 GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle 377  
QY 182 CTACTAGATCGTCCCGAGACTCGCCCTACAGATGAGTGTGATCTCACCTGAATATTATATC 241

Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
QY 242 CAGTAGAAGCATGCTATGAAAGGACAGACCTATTGAGTCAGTGCCTGATCTGCG 301  
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACCCCATAATTATCTATCAAGAAACACAGCATCTGGTAATATCCATGACATCTTT 361  
Db 418 ThrProLeuIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 437  
QY 362 CATGTTTTCCTCCCAAGTCACAGAGAAATGAGTTTATTTTGGCTCTGAAATGCAA 421  
Db 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGGTTTCGTCATTTATACAAATATCATCTATTTTAAAGGAAAGCAATATAACGA 481  
Db 458 ThrGlyPheArgHisLeuIleIleIleIleIleIleIleIleIleIleIleIleIleIle 477  
QY 482 TCCAGTGGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA 541  
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497  
QY 542 ATTCACAGTGGTGAATGGAGGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAA 601  
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTCAGAGGCTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAC 661  
Db 518 ValArgArgLeuValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 537  
QY 662 GTAGTCAGTGTACGTAATCTGGAGGTGACAGGCTGACTGACCGTGGCTACTCACAT 721  
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
QY 722 TCTTGCTGATCAGTCACACTGTGACTCTTTATAGTAAGTATAGTAAACAGAGAGAT 781  
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerIleIleIleIleIleIleIle 577  
QY 782 CCACACTGNGTCCCTTTTCAAGCTATCAAGCTGACCTGAGATCAACCACTTGCAGAA 841  
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThr 597  
QY 842 RAGGAATTTGGGCCACCATTTTGGATTCAGAGTCTCTCTTCTGACTATATCTCTCCA 901  
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
QY 902 GAAATTTCTCTTTGAAAGTACTGGAATTTACATGATGGGATGCTCTCAAGCT 961  
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
QY 962 CATGATCTACAGCTCGGAAAGAAATATCTACTGCTGTCTATATATGATGGTGGTCTCAG 1021  
Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGln 657  
QY 1021 ----- 1021  
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValIleTyrPheArgLeuAsnThrLeuAla 677  
QY 1021 ----- 1021  
Db 678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
QY 1021 ----- 1021  
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
QY 1021 ----- 1021  
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021

Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTCTATTCTCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAA 1081  
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777  
QY 1082 CGTTATATGGGTCAACCTTCCAGCAATGAAAGGCTATTACTTAGGATCTGTGGCCCATG 1141  
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797  
QY 1142 CAAGCAGAAAGTTCCCTCTCGAACCAATCCGTTTACTGCTTACATGGTTTCTCTGGAT 1201  
Db 798 GluAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAsp 817  
QY 1202 GAGAAATGTCATATTTCCACATACCAAGTATATTAATCTAGTGTGAGGCTGGAAAG 1261  
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
QY 1262 CCATATGATTTACAGATCTATCTCAGGAGAGACACAGATAAGAGTTCTCTGAATCGGA 1321  
Db 838 ProTyrAspLeuGlnIleTyrProGlnArgHisSerIleArgValProGluSerGly 857  
QY 1322 GAACATTATGAATGCTCATCTTTTGGCTACCTTCAAGAAACCTTGGATCAGCTATTGCT 1381  
Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluLeuLeuGlySerArgIleAla 877  
QY 1382 GCTCTAAGAGTGATA 1396  
Db 878 AlaLeuLysValIle 882  
RESULT 4  
US-09-976-674-7  
; Sequence 7, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR FILING DATE: US 60/240,117  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-7  
Alignment Scores:  
Pred. No.: 3,61e-221 Length: 690  
Score: 1957.50 Matches: 372  
Percent Similarity: 99.5% Conservative: 1  
Best Local Similarity: 99.2% Mismatches: 0  
Query Match: 66.2% Indels: 2  
DB: Gaps: 1  
US-10-825-632-6 (1-1669) x US-09-976-674-7 (1-690)  
QY 2 ACAGGTACAGCAATCTTAAGTCACATTTTAAGATGTCAGAAATAATGATTGATCTGAA 61  
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337  
QY 62 GGAAGATCATAGATGTCATAGATAGAGAACTAATTAACCTTTTCTAGATTCTATTGAA 121  
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357  
QY 122 GGAGTTGAATATATTGCGAGAGTGGATGACTCTCGAGGAAATATATCTGGTCCATC 181

358 GlyValGluTyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaThrSerIle 377  
182 CTACTAGATCGCTCCAGACCTCGCTACAGATGAGTGTGATCTCACCTGAATATTATC 241  
378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
242 CCAGTGAAGATGATGTTAAGAACGACAGACTCATGTCAGTCCCTGATCTGTCG 301  
398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
302 ACCCATTAATCTATGAGAAACAACAGACATCTGATAAATATCCATGACATCTTT 361  
418 ThrProLeuIleIleTyrGluGluThrThrAspIleThrIleAsnIleHisAspIlePhe 437  
362 CATGTTTTTCCCAAAAGTCACGAGAGGAAATAGTATTATTTTGGCTCTGAATGAAA 421  
438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457  
422 ACAGTTTCCGTCATTTATACAAATATCATCTATTATTAAGGAAGCAATATAACGA 481  
458 ThrGlyPheArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
482 TCACAGTGGTGGCTGCTGCTCAAGTGTCTCAAGTGTCTCAAGAGAGAGATAGCA 541  
478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497  
542 ATTACCAAGTGGTGAATGGGAAAGTCTTCTGCGGATGATCTTAATATCCAAAGTGA 601  
498 IleThrSerGlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
602 GTCAGAGGCTGGTATATTTGAAGCACCAGACTCCCTTTAGACCATCCTGTAC 661  
518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537  
662 GTACTGATTAACGTAATCTGGAGAGGTGACAGGCTGACAGCTGACCGTGGTACTCAT 721  
538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
722 TCTTGCTGCATCAGCAGCTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAGAA 781  
558 SerCysCysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsn 577  
782 CACACTGTGTGCTCCCTTTACAGCTATCAAGTCTCGAAGATCACCCACTTGCAAAACA 841  
578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
842 AAGGAATTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTGACTATATCTCTCCA 901  
598 LysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
902 GAAATTTTCTCTTTTGAAGTACTACTGGATTTACATGTTGATGGGATGCTCTACAGCCT 961  
618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
962 CATGATCTACAGCTCGGAAGAAATATCTACTGCTGCTGCTCATATATGCTGCTCTCAG 1021  
638 HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGly---Arg 656  
1022 GTTGTATTGCTGGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAA 1081  
657 -LeuLeuLeuLeuGlyProGlnSerLeuCysGlySerSerMetIleGlnAspThrArgAs 676  
1082 CDTTATATGGGTACCCCTGACCAGAAATGAACAGGGCTATTACT 1124  
676 nValIleThrValThrLeuThrArgMetAsnArgAlaIleThr 690

APPLICANT: Riviere, Pierre  
APPLICANT: Junien, Jean-Louis  
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
FILE REFERENCE: 70669  
CURRENT APPLICATION NUMBER: US/09/976,674  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: US 60/240,117  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 19  
LENGTH: 658  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-976-674-19

Alignment Scores:  
Pred. No.: 1,5e-203 Length: 658  
Score: 1808.00 Matches: 338  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 61.1% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-6 (1-1669) x US-09-976-674-19 (1-658)

QY 2 ACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATGTCAAGAAATATGATGCTGAA 61  
DB 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337  
QY 62 GGAAGGATCATPAGATGTCTAGATAAGAACTAATTAACCTTTTGAGATTTCTATTGAA 121  
DB 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357  
QY 122 GGAGTTGAATATATGTCAGAGCTGGAGTGGACCTCTGAGGAAATATCTTGGTCCATC 181  
DB 358 GlyValGluTyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaThrSerIle 377  
QY 182 CTACTAGATGCTCCAGACTCGCTACAGATAGTGTGATCTCATCTCAATATTATTTATC 241  
DB 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
QY 242 CCAGTAGAAGATGATGTTATGAAAGCAGAGACTCATGTCAGTCCCTGATCTGTCG 301  
DB 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 AGCCCACTAATTTATCTATGAAAGAACAGACATCTGGATAAATATCCATGACATCTTT 361  
DB 418 ThrProLeuIleIleTyrGluGluThrThrAspIleThrIleAsnIleHisAspIlePhe 437  
QY 362 CATGTTTTTCCCAAAAGTCACGAGAGGAAATAGAGTTTATTTTGCCTCTGAATGAAA 421  
DB 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGTTTCCGTCATTTATACAAATATCATCTATTATTAAGGAAGCAATATAACGA 481  
DB 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
QY 482 TCCAGTGGTGGCTGCTGCTCCAGTGTCTCAAGTGTCTCATCAAGAGAGAGATAGCA 541  
DB 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497  
QY 542 ATTACCAAGTGGTGAATGGGAAAGTCTTCTGCGGATGATCTTAATATCCAAAGTGA 601  
DB 498 IleThrSerGlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTCAAGAGGCTGGTATATTTGAAGCACCAGACTCCCTTTAGACCATCCTGTAC 661  
DB 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537  
QY 662 GTAGTCAGTTAGTAAATCTCGAGAGGTGCAAGGCTGACTGACCGTGGTACTCATCAT 721  
DB 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557





```
Pred. No.: 5.84e-172 Length: 613
Score: 1540.00 Matches: 290
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 52.1% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-6 (1-1669) x US-09-976-674-21 (1-613)

QY 2 ACAGGTACAGCAAACTCTAAAGTCACATTTTAAGATCTCAGAAATAATGATGATGCTGAA 61
Db ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATCTTATTTGAA 121
Db GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 CGAGTTGGAATATATGTCAGAGCTGGATGCTCTGAGGGAATAATGCTGGTCCATC 181
Db GlyValGluTyrIleAlaAaGAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGATCGCTACAGATAGTGTGATCTCACCTGAAATTTATTC 241
Db LeuLeuAspArgSerGlnThrArgLeuGlnIleValIleSerProGluLeuPheIle 397
QY 242 CCAGTAGAAGATGATGTTATGGAAGGACAGACTCATTGAGTCAGTCGCTGATCTGTG 301
Db ProValGluAspAspValMetGluAtgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 AGCCACATAATTCATATGAGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 361
Db ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTTCCCAAGTCACGAGAGGAATTTGATTTATTTTTCCTCTGAATGCAAA 421
Db HisValPheProGlnSerHisGluGluGluLeuPheIlePheAlaSerGluCysLys 457
QY 422 ACAGGTTTCCTCATTTATACAAATTTACATCTATTATTTAAAGGAAGCAAAATATAAACA 481
Db ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
QY 482 TCACGTGTGGGCTGCTGCTCCAGTGATTCATGATTCCTATCAAGAGAGATAGCA 541
Db SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
QY 542 ATTACCACTGTGTAATCGGAAGTCTTGGCCGSCATGATCTAATATCAAGTTGATGAA 601
Db IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTCAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGACATCACCTGTAC 661
Db ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
QY 662 GTAGTCATTACGTAATCTCGAGAGGTGACAAAGCTCACTGACCGTGGCTACTACAT 721
Db ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTTGTCTCATCTACGACTGCTGACTCTTATTAAGTAAGTATAGTAAACAGAGAGAT 781
Db SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CACACTGTGTGCTCCCTTTACAGCTATCAAGTCTCGAAGATGACCCCACTTGCAAAACA 841
Db ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
QY 842 AAGGAATTTGGCCACCATTTTGATTC 871
Db LysGluPheTrpAlaThrIleLeuAspSer 607
```

## RESULT 8

US-09-976-674-3

; Sequence 3, Application US/09976674

; Patent No. 6844180

```
GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-3

Alignment Scores:
Pred. No.: 1.72e-166 Length: 863
Score: 1494.50 Matches: 281
Percent Similarity: 62.9% Conservative: 69
Best Local Similarity: 50.5% Mismatches: 105
Query Match: 50.5% Indels: 101
DB: 2 Gaps: 2

US-10-825-632-6 (1-1669) x US-09-976-674-3 (1-863)
```

```
QY 2 ACAGGTACAGCAAACTCTAAAGTCACATTTTAAGATCTCAGAAATAATGATGATGCTGAA 61
Db ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 327
QY 62 GGAAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATCTTATTTGAA 121
Db GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 347
QY 122 GGAGTTGAATATATTTGCAGAGCTGGATGCTCTCTGAGGAAATAATGCTTTGGTCCATC 181
Db LysValGluTyrIleAlaAaGAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMet 367
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTC 241
Db PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIle 387
QY 242 CCAGTAGAAGATGATGTTATGGAAGGCACAGACTCATTGAGTCAGTCGCTGATCTGTG 301
Db ProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal 407
QY 302 ACGCCACTAATTATCTATGAAGAAACACACAGACATCTGGATAAATATCCATGACATCTTT 361
Db GlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePhe 427
QY 362 CATGTTTTTCCCAAGCTCAC---GAAGAGAAATTTGACTTTATTTTGGCTCTGCAATGC 418
Db TrpProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys 447
QY 419 AAAACAGGTTTCCGTCATTTATACAAATTTACATCTATTATTAAGAAAGAAACAAATATAA 478
Db LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 467
QY 479 CGATCCAGTGTGGGCTGCTCTCCAAAGTATTTCAAGTGTCTCTATCAAGAGGAGAGATA 538
Db TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGluIle 487
QY 539 GCAATTAACCATGTGTGAATCGGAAGTCTTTGGCCGCATGATCTAATATCAAGTTGAT 598
Db AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAsn 507
QY 599 GAACTCAGAAGCTGGTATATTTTGAAGGCACCAAGACTCCCTCTAGAGCATCACCTG 658
Db GluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeu 527
```

Qy	659	TACGTAGTCAGTTACCTAAATCTCGAGAGGTGACAAAGCTGACTGACCGTGCTACTCA	718	;	APPLICANT: Oi, Steve
Db	528	TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrProGlyPheSer	547	;	APPLICANT: Akineanya, Karen
Qy	719	CATTCTTGTCATCAGTCAGCACGTGTGACTTCTTTATAGTAAGTATAGTAACCAAG	778	;	APPLICANT: Riviere, Pierre
Db	548	HisSerCysSerMetSerGlnAenPheAspMetPheValSerHisTyrSerValSer	567	;	APPLICANT: Junien, Jean-Louis
Qy	779	AATCCACACATGTGTCTTACAGCTATCAAGCTCTCAAGTCTGAAGTACCAACCTTGCA	838	;	FILE REFERENCE: 70669
Db	568	ThrProProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys	587	;	TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
Qy	839	ACAAAGAAATTTGGCCACCAATTTGGATTTCAGAGCTCTCTCTCTGACTATCTCT	898	;	FILE REFERENCE: 70669
Db	588	GlnProArgPheTrpAlaSerMetMetGluAlaHisSerCysProProAspTyrValPro	607	;	CURRENT APPLICATION NUMBER: US/09/976,674
Qy	899	CCGAAATTTCTCTTTCAAGTACTACTGATTTACATTTGATGATGGATGCTCAAG	958	;	CURRENT FILING DATE: 2001-10-12
Db	608	ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys	627	;	PRIOR APPLICATION NUMBER: US 60/240,117
Qy	959	CCTCATGACTCAGCTCGAAGAAATATCTACTGCTGCTTCATATATGCTGCTCT	1018	;	PRIOR FILING DATE: 2000-10-12
Db	628	ProHisAlaLeuGlnProGlyLysHisProHisValLeuPheValTyrGlyGlyPro	647	;	NUMBER OF SEQ ID NOS: 61
Qy	1019	CAG-----	1021	;	SOFTWARE: Patentin version 3.1
Db	648	GlnValGlnLeuValAenAenSerPheLysGlyIleLysTyrLeuArgLeuAenThrLeu	667	;	SEQ ID NO 23
Qy	1021	-----	1021	;	LENGTH: 892
Db	668	AlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeu	687	;	TYPE: PRT
Qy	1021	-----	1021	;	ORGANISM: Homo sapiens
Db	688	ArgPheGluGlyAlaLeuLysAenGlnMetGlyGlnValGluIleGluAenGlnValGlu	707	;	US-09-976-674-23
Qy	1021	-----	1021	;	Alignment Scores:
Db	708	GlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis	727	;	Pred. No.: 1,75e-166
Qy	1021	-----	1021	;	Score: 1494.50
Db	728	GlyTrpSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe	747	;	Matches: 281
Qy	1022	---GTTGCTATTCTGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACAG	1078	;	Percent Similarity: 62.9%
Db	748	LysValAlaIleAlaGlyAlaProValThrValTyrMetAlaTyrAspThrGlyTyrThr	767	;	Conservative: 69
Qy	1079	GAAGTTATATGGTCACCTCAGCAGATGAACAGGCTATTACTTGGATCTGTGGCC	1138	;	Best Local Similarity: 50.5%
Db	768	GluArgTyrMetAspValProGluAenAenGlnHisGlyTyrGluAlaGlySerValAla	787	;	Query Match: 50.5%
Qy	1139	ATGCAAGCAGAAAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGCTTCTG	1198	;	DB: 2
Db	788	LeuHisValGluLysLeuProAenGluProAenArgLeuLeuIleLeuHisGlyPheLeu	807	;	US-10-825-632-6 (1-1669) x US-09-976-674-23 (1-892)
Qy	1199	GATGAGATGTCATTTGCACATACACAGTATATTAAGTATTTTATTTTTCCTG	1258	;	Qy 2 ACAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTCAGAAATTAATGATGCTGAA
Db	808	AspGluAenValHisPhePheHisThrAenPheLeuValSerGlnLeuIleAeGalaGly	827	;	Db 337 ThrGlySerLysAenProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln
Qy	1259	AAGCATATGATTTAAGATATATCTCAGAGAGACACAGATTAAGATGCTCTGAATCG	1318	;	Qy 62 GGAAGATCATGATGTCATAGATAGGAACCTTAACTTCAACCTTTTGAGTATCTATTGAA
Db	828	LysProTyrGlnLeuGlnIleTyrProAenGluArgHisSerIleAeGlyProGluSer	847	;	Db 357 GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro
Qy	1319	GGAGAACATTATCAATGCTCTTTTGGCACTACCTCTCAAGAAACCTT 1366		;	Qy 122 GGAATTGAATATATCCAGAGCTGATGAGTCTCTCGAGGAAAATATGCTGTCATC
Db	848	GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 863		;	Db 377 LysValGluTyrIleAlaArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMet
RESULT 9				;	Qy 182 CTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACCTGAAATATTATC
US-09-976-674-23				;	Db 397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIle
; Sequence 23, Application US/09976674				;	Qy 242 CCAGTAGAATGATGTTATGGAAGGAGGAGACTCATTGAGTCAGTCCCTGATCTGTG
; Patent No. 6844180				;	Db 417 ProSerThrGluAenGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAenVal
; GENERAL INFORMATION:				;	Qy 302 ACGCCACTAATATTCTATCAAGAAACACAGACATCTGATAAATATCCATGACATCTTT
				;	Db 437 GlnProTyrValValTyrGluGluValThrAenValTyrIleAenValHisAspIlePhe
				;	Qy 362 CATGTTTTCCTCCAAAGTCAC--GAAGAGGAAATTTAGTTTATTTTTCCTCTGAATGC
				;	Db 457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAenGluCys
				;	Qy 419 AAACAGGTTTCCTCATTTATACAAATTTACATCTTATTTAAAGGAAAGCAATATAAA
				;	Db 477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp
				;	Qy 479 CGATCCAGTGGTGGCTGCTCTCCAGTGTATTTCAAGTATTTCAAGTCTCTCAAGAGAGATA
				;	Db 497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysLeuIle
				;	Qy 539 GCAATACAGTGGTGAATGGAGGTTCTTGGCCGCGATGATCTTAATATCAAGTTGAT
				;	Db 517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAen
				;	Qy 599 GAAGTCAGAGGCTGGTGTATTTTGAAGGCACAAAGACTCCCTTTAGACATCACTG
				;	Db 537 GluGluThrLysLeuValTyrPheGlnGlnThrLysAspThrProLeuGluHisLeu
				;	Qy 659 TACGTAGTCAGTTACGTAATAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCA

```
Db      557 TyrValValSerTyrGluAlaAlaGlyGluLeuValArgLeuThrThrProGlyPheSer 576
Qy      719 CATTCTTCTGCATCAGTCAGCAGTCGTGACTTCTTTATAGTAAAGTATAGTAAACAGAG 778
Db      577 HisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerValSer 596
Qy      779 AATCACAACATGTGTCTCTTTTACAAGCTATCAAGTCTCAAGTCTCAAGTCAAGTCAAGTCAAG 838
Db      597 ThrProCysValHisValTyrLysLeuSerGlyProAspAspPheLeuHisLys 616
Qy      839 ACAAGGAATTTGGCCGCCCATTTGGATTTCAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 898
Db      617 GlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProAspPyrValPro 636
Qy      899 CAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTTGATGGATGGATGGATGGATGGATGG 958
Db      637 ProGluLeuPheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys 656
Qy      959 CTTCTATGATCTACAGCCCTGGAAGAAATATCTCTACTGTCTGTCTGTCTGTCTGTCTGTCTGT 1018
Db      657 ProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGlyPro 676
Qy      1019 CAG----- 1021
Db      677 GlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeu 696
Qy      1021 ----- 1021
Db      697 AlaSerLeuGlyTyrAlaValValValLeuAspGlyArgGlySerCysGlnArgGlyLeu 716
Qy      1021 ----- 1021
Db      717 ArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGlu 736
Qy      1021 ----- 1021
Db      737 GlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis 756
Qy      1021 ----- 1021
Db      757 GlyTrpSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe 776
Qy      1022 ----GTTCGCTATTGCTGGGCCCGCCAGTCCTCTGTGGATCTTCTATGATACAGGATACAG 1078
Db      777 LysValAlaIleAlaGlyAlaProValThrValThrValThrMetAlaTyrAspThrGlyTyrThr 796
Qy      1079 GAACGTTATATGGGTCAACCTGACAGCAATGAACAGGGCTATTACTTTAGGATCTGTGGCC 1138
Db      797 GluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAla 816
Qy      1139 ATGCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCCTG 1198
Db      817 LeuHisValGlyLysLeuProAsnGluProAsnArgLeuLeuLeuHisGlyPheLeu 836
Qy      1199 GATGCAAGTCCATTTGACATACACAGTATATTACTAGTATTTTGTAGGCGGTGGA 1258
Db      837 AspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGly 856
Qy      1259 AAGCCATATGATTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTCTCTGAATCG 1318
Db      857 LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGluSer 876
Qy      1319 GAGAACATATGAACTGATCTTTTGCATCTTTTGCATCTTTTCAAGAAACCTT 1366
Db      877 GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 892
```

RESULT 10

US-09-976-674-27

; Sequence 27, Application US/09976674

; Patent No. 6844180

; GENERAL INFORMATION:

; APPLICANT: Qi, Steve

```
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 27
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-27
```

```
Alignment Scores:      1.75e-166      Length:      892
Pred. No.:             1494.50         Matches:     281
Score:                 62.9%           Conservative: 69
Percent Similarity:    50.5%           Mismatches:  105
Best Local Similarity: 50.5%           Indels:       101
Query Match:           2               Gaps:         2
DB:
```

US-10-825-632-6 (1-1669) x US-09-976-674-27 (1-892)

```
Qy      2 ACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGTCTGAA 61
Db      337 ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 356
Qy      62 GGAAGGATCATAGATGTCTATAGATAAGAACTAATCAACCTTTTGAGATCTATTGTAA 121
Db      357 GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 376
Qy      122 GGAATTGATATATTCAGAGAGCTGGATGGACTCTCGAGGAAATATGCTGTGTCATC 181
Db      377 LysValGluTyrIleAlaArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMet 396
Qy      182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTAGTTGTATCTCACCTGAATTTATTATC 241
Db      397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIle 416
Qy      242 CCAGTAAAGATGATGTTTATGAAAGGAGAGAGACTCATGTAGTCAAGTCAAGTCAAGTCA 301
Db      417 ProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal 436
Qy      302 ACCCACTAATTATCTATGAAGAAACAAGACATCTGGATAAATATCCATGATCTTT 361
Db      437 GlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePhe 456
Qy      362 CATGTTTTTCCCAAGTCAC--GAAGAGAAATTTAGTTTATTTTGGCTCTGAATGC 418
Db      457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys 476
Qy      419 AAAACAGAGTTCCGTCATTATATCAAAATACATCTATTATTTAAAGAAAGCAAAATATAA 478
Db      477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 496
Qy      479 CGATCCAGTGGTGGGCTGCTCTCCAAGTGNATTCAAGTGTCTCAAGAGAGAGAGATA 538
Db      497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGluIle 516
Qy      539 GCAATACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATATCCAAAGTTGAT 598
Db      517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysLeuTrpValAsn 536
Qy      599 GAAAGTCAGAAAGCGCTGATATATTTGAAGGACCAAGACTCCCTTTTAGACATCACTG 658
Db      537 GluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeu 556
Qy      659 TAGGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTGACTGACCGGTGCTACTCA 718
```

```
Db 557 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSer 576
Qy 719 CATTTCTGTCAGTACAGTCTGACTCTTTTATAAGTAAGTATAGTAACAGAG 778
Db 577 HisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerValSer 596
Qy 779 AATCCACACTGTGTGCTTACAGCTCTCAAGCTATCAAGTCTGAAGATGACCCCACTTCACA 838
Db 597 ThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys 616
Qy 839 ACAAGAAATTTGGCCACCATTTTGGATTTCAGAGCTCTCTTCTGCTACTACTCT 898
Db 617 GlnProArgPheTrpAlaSerMetMetGluAlaLaserCysProProAspTyrValPro 636
Qy 899 CCAAGAAATTTCTCTTTTCAAGTACTACTGATTTCATTTGATGGATGCTCTCAAG 958
Db 637 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys 656
Qy 959 CCTCATGATCAGCTGGAAGAAATATCTACTGTCTGCTTCATATATGTGTCTCT 1018
Db 657 ProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyPro 676
Qy 1019 CAG----- 1021
Db 677 GlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeu 696
Qy 1021 ----- 1021
Db 697 AlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeu 716
Qy 1021 ----- 1021
Db 717 ArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGlu 736
Qy 1021 ----- 1021
Db 737 GlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis 756
Qy 1021 ----- 1021
Db 757 GlyTrpSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe 776
Qy 1022 ---GTTGCTATTGCTGGGCCCCAGTCACTCTGTTGGATCTTATCATACAGGATACAGG 1078
Db 777 LysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 796
Qy 1079 GAACGTTATGGGTCACTGACCAAGTCAACAGGCTTACTTAGATCTCTGGCC 1138
Db 797 GluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAla 816
Qy 1139 ATGCAAGCAGAAAGTTCCTCTCAACCAATCGTTACTGCTCTTACATGTTTCTG 1198
Db 817 LeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleLeuHisGlyPheLeu 836
Qy 1199 GATGAGATGTCCATTGTCACATACCATATATTTACTGAGTCTTCTAGGCTCGA 1258
Db 837 AspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGly 856
Qy 1259 AAGCCATATGATTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTCTCTCAATCG 1318
Db 857 LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGluSer 876
Qy 1319 GGAGAACATTATGACATCTTTTTCACATCTTTCCTCACTCTTCAAGAAACCTT 1366
Db 877 GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGlyTrpLeu 892
```

## RESULT 11

```
US-09-976-674-33
; Sequence 33, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
```

```
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 20649
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 33
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-33
```

## Alignment Scores:

```
Pred. No.: 1,2e-160 Length: 879
Score: 1445.00 Matches: 274
Percent Similarity: 61.3% Conservative: 67
Best Local Similarity: 49.3% Mismatches: 101
Query Match: 48.9% Indels: 114
DB: 2 Gaps: 3
```

US-10-825-632-6 (1-1669) x US-09-976-674-33 (1-879)

```
Qy 2 ACAGGTACAGCAAAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGAA 61
Db 337 ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 356
Qy 62 GGAAGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGGAGATTTCTATTGAA 121
Db 357 GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 376
Qy 122 GGAATTTGATATATTCAGCAGAGCTGGAGTCTCTCAGGAGAAATATGTTGGTCCATC 181
Db 377 LysValGluTyrIleAlaArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMet 396
Qy 182 CTACTAGATCGTCCCACTCGCTACAGATAGTGTGATCTCACTCAATGATTTATTATC 241
Db 397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIle 416
Qy 242 CAGGTAGAATGATGTTATGGAAGGACAGACTCAATGAGTCAGTCCTGATTTCTGTG 301
Db 417 ProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal 436
Qy 302 AGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATPCCATGACATCTTT 361
Db 437 GlnProTyrValValTyrGluValThrAsnValTrpIleAsnValHisAspIlePhe 456
Qy 362 CAGTTTTTCCCAAGTCAC---GAAGAGAAATTCAGTTTATTTTTCCTCTGAATGC 418
Db 457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys 476
Qy 419 AAAACAGGTTTCCGTCATTATACAAAATTCATCTATTATTTTAAAGGAAAGCAAAATATAA 478
Db 477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 496
Qy 479 CGATCCAGTGTGGCTGCTCTCCCAAGTGATTTCAGGTGCTCTATCAAGAGAGAGATA 538
Db 497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluIle 516
Qy 539 GCAATTACAGTGTGAATGGAAAGTTCTTGGCGGCGATCGATCTAATATCCAAGTTGAT 598
Db 517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySer----- 531
Qy 599 GAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTCTTTAGAGATCACCTG 658
Db 532 -----LysGlyThrLysAspThrProLeuGluHisLeu 543
Qy 659 TAGGTAGTCACTTACGTAATAATCTGGAGAGGTCACAGGCTGACCTACCGTGGCTACTCA 718
Db 544 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSer 563
```

QY 719 CATTCTTCTGCATCAGTCAGCTGCTGACTCTCTTTATAGTAAGTATAGTAACCAAG 778  
Db 564 HisSerCysSerMetSerGluAsnPheAspMetPheValSerHisTyrSerValSer 583  
QY 779 AATCCACACTGTGTCTCTTTACAAGCTATCAAGTCTGAAGATGACCAACTTGCAAA 838  
Db 584 ThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys 603  
QY 839 ACAAAGGAATTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATCTCT 898  
Db 604 GlnProArgPheTrpAlaSerMetMetGluAlaSerCysProProAspTyrValPro 623  
QY 899 CCAGAAATTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAAG 958  
Db 624 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys 643  
QY 959 CCTCATGATCTACACCTCGGAAGAAATATCTACTGTCTGTTCATATATGTTGCTCT 1018  
Db 644 ProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyPro 663  
QY 1019 CAG----- 1021  
Db 664 GlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeu 683  
QY 1021 ----- 1021  
Db 684 AlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeu 703  
QY 1021 ----- 1021  
Db 704 ArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGlu 723  
QY 1021 ----- 1021  
Db 724 GlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIleHis 743  
QY 1021 ----- 1021  
Db 744 GlyTrpSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPhe 763  
QY 1022 ---GTTGCTATTGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAG 1078  
Db 764 LysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 783  
QY 1079 GAACGTTATATGGTCACCTCAGCAGAAATGAAAGAGGCTTATCTAGGATCTGGCC 1138  
Db 784 GluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAla 803  
QY 1139 ATGCAAGCAGAAAGTTCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTG 1198  
Db 804 LeuHisValGluLysLeuProAsnGluProAsnArgLeuIleLeuHisGlyPheLeu 823  
QY 1199 GATGAGAATCTCCATTTTGCACATACAGTATATCTAGTCTTTTGTAGCGGCTGGA 1258  
Db 824 AspGluAsnValHisPheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGly 843  
QY 1259 AAGCCATATGATTACAGATCTATCTCTCAGGAGAGACACACATAAGAGTTCTCGAATCG 1318  
Db 844 LysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGluSer 863  
QY 1319 GAGAACATATGAACTGATCTTTTGCATCTACTCTTCAAGAAACCTT 1366  
Db 864 GlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 879

## RESULT 12

US-09-976-674-35  
; Sequence 35, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 35  
; LENGTH: 879  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-35  
Alignment Scores:  
Pred. No.: 1-2e-160 Length: 879  
Score: 1445.00 Matches: 274  
Percent Similarity: 61.3% Conservatives: 67  
Best Local Similarity: 49.3% Mismatches: 101  
Query Match: 48.9% Indels: 114  
DB: 2 Gaps: 3

US-10-825-632-6 (1-1669) x US-09-976-674-35 (1-879)

QY 2 ACAGGTACAGCAAAATCCTAAAGTCATCTTTTAAGATGTCAGAAATAATGATGATGCTGAA 61  
Db 337 ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 356  
QY 62 GGAAGATCATAGATGTCATAGATGAAGCACTAATTAACCTTTTGAGATCTTATTTGAA 121  
Db 357 GlyLysIleValSerThrGlnLysGluLeuValGlnProPheSerSerLeuPhePro 376  
QY 122 GGAGTTGAATATATGTCAGAGCTGGATGGACTCTCTGAGGGAAATATGCTTGCCATC 181  
Db 377 LysValGluTyrIleAlaArgAlaGlyTyrThrArgAspGlyLysTyrAlaTrpAlaMet 396  
QY 182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGTGATCTCACCTGAATTTATTTATC 241  
Db 397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuProAlaLeuPheIle 416  
QY 242 CCGATGAGATGATGTTTATGGAAGGAGAGACTCATTGAGTCAGTCCCTGATTCGTG 301  
Db 417 ProSerThrGluAsnGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal 436  
QY 302 ACCGCAATATATCTATCAAGAAACAAACAGACATCTCGATAAATATCCATGACATCTTT 361  
Db 437 GlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePhe 456  
QY 362 CATGTTTTTCCCAAGTCCAC---GAAGAGGAATTTGATTTTATTTTGCCTCTGAATGC 418  
Db 457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys 476  
QY 419 AAAACAGGTTTCGTCATTTATACAAATATACATCTATTATAAGGAAGCAAAATATAA 478  
Db 477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 496  
QY 479 CGATCCAGTGGTGGGCTGCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGAGATA 538  
Db 497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGluIle 516  
QY 539 GCAATACCAAGTGGTGAATGGGAAGTCTTGGCCGCGCATGGATCTTAATATCAAGTTGAT 598  
Db 517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySer----- 531  
QY 599 GAAGTCAGAAGGCTGGTGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTG 658  
Db 532 -----LysGlyThrLysAspThrProLeuGluHisLeu 543  
QY 659 TACGTAGTCAAGTACGTAAATCCTCGAGAGAGGTGACAGGCTGACTGACCGTGGCTACTCA 718  
Db 544 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSer 563







Db 337 ThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGln 356  
QY 62 GGAAGGATCATAGATGTCTAGATAAGGAACATAATCAACCTTTTGAGATTCTATTGAA 121  
Db 357 GlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSerSerLeuPhePro 376  
QY 122 GGAAGTTGAATATATCCAGAGCTGGATGGACTCTGAGGAGAAATATGCTTGTCATC 181  
Db 377 LysValGluTyrIleAlaArgAlaGlyThrArgAspGlyLysTyrIleAlaMet 396  
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTATC 241  
Db 397 PheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuProAlaLeuPheIle 416  
QY 242 CCAGTAGAGATGATGTTATGAAAGGCAGAGACTCATTTGAGTCAGTCCCTGATTCTGTG 301  
Db 417 ProSerThrGluAsnGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnVal 436  
QY 302 ACCCCACTAATTATCTATGAAGAACACAGACATCTGGATAATATCCATGACATCTTT 361  
Db 437 GlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIlePhe 456  
QY 362 CATGTTTCCCAAGTCAC---GAAGAGAAATGAGTTTATTTTGGCTCTGAATGC 418  
Db 457 TyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCys 476  
QY 419 AAACAGGTTCCGTCATTATACAAATTACATCTATTATAAGGAAAGCAAAATATAA 478  
Db 477 LysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAsp 496  
QY 479 CGATCCAGTGTGGGCTCGCTCCAGTCATTTCAAGTGTCTCTATCAAGAGGAGGATA 538  
Db 497 TrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluLeu 516  
QY 539 GCATTTACCAAGTGTGAATGGAAGTCTTGGCGGCATGGATCTAATATCCAAAGTTGAT 598  
Db 517 AlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpValAsn 536  
QY 599 GAAGTCAGAGCTGGTATATTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTG 658  
Db 537 GluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeu 556  
QY 659 TAGCTAGTCACTAGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCA 718  
Db 557 TyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSer 576  
QY 719 CATTCCTGCTCAGTCAGTCAGTGTGACTCTTTTAAAGTAAGTATAGTAAACCAAG 778  
Db 577 HisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerValSer 596  
QY 779 AATCCACTGTGTCCCTTTCAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAA 838  
Db 597 ThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHisLys 616  
QY 839 ACAAGGAATTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCCT 898  
Db 617 GlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrValPro 636  
QY 899 CCAGAAATTTCTTTTGAAGTACTACTGGATTTCATTTGATGGATGCTCTACAAG 958  
Db 637 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLys 656  
QY 959 CCTCATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGCTCATATATGGTGGCT 1018  
Db 657 ProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyPro 676  
QY 1019 CAGGTTGCTATT 1030  
Db 677 GlnValGlnLeu 680

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 02:43:09 ; Search time 47.8004 Seconds  
(without alignments)  
2917.786 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 2957  
Sequence: 1 aacaggtacagcaatccta.....aaaaaaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSSWEB spool/US10825632/runat 01052006 105957 3449/app\_query.fasta\_1  
-DB=Published Applications AA Main -QFMT=fastan -SUFFIX=n2p.rapbm  
-MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1  
-MATRIX=blosum62 -TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p  
-USBR=US10825632 @CGN 1.1 503 @runat 01052006 105957 3449 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=10 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : Published Applications AA Main:

1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
4: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	2482	83.9	465	4	US-10-825-632-5	Sequence 5, Appli
2	2422	81.9	882	3	US-09-976-674-1	Sequence 1, Appli
3	2422	81.9	882	4	US-10-054-776-2	Sequence 2, Appli
4	2422	81.9	882	4	US-10-170-789-38	Sequence 38, Appli
5	2422	81.9	882	4	US-10-311-035-9	Sequence 9, Appli
6	2422	81.9	882	4	US-10-072-012-622	Sequence 622, App
7	2422	81.9	882	4	US-10-415-122-6	Sequence 6, Appli
8	2422	81.9	882	4	US-10-825-632-1	Sequence 1, Appli
9	2422	81.9	882	5	US-10-982-512-1	Sequence 1, Appli
10	2326	78.7	883	4	US-10-072-012-621	Sequence 621, App
11	2244.5	75.9	580	4	US-10-275-505-2	Sequence 2, Appli

12	2244.5	75.9	580	6	US-11-140-224-2	Sequence 2, Appli
13	1957.5	66.2	690	3	US-09-976-674-7	Sequence 7, Appli
14	1957.5	66.2	690	3	US-10-982-512-7	Sequence 7, Appli
15	1808	61.1	658	3	US-09-976-674-19	Sequence 19, Appli
16	1808	61.1	658	5	US-10-982-512-19	Sequence 19, Appli
17	1808	61.1	661	3	US-09-976-674-11	Sequence 11, Appli
18	1808	61.1	661	5	US-10-982-512-11	Sequence 11, Appli
19	1540	52.1	613	3	US-09-976-674-21	Sequence 21, Appli
20	1540	52.1	613	5	US-10-982-512-21	Sequence 21, Appli
21	1494.5	50.5	689	4	US-10-072-012-620	Sequence 7, Appli
22	1494.5	50.5	830	4	US-10-415-122-7	Sequence 3, Appli
23	1494.5	50.5	863	3	US-09-976-674-3	Sequence 619, App
24	1494.5	50.5	863	4	US-10-072-012-619	Sequence 3, Appli
25	1494.5	50.5	863	5	US-10-982-512-3	Sequence 23, Appli
26	1494.5	50.5	892	3	US-09-976-674-23	Sequence 27, Appli
27	1494.5	50.5	892	3	US-09-976-674-27	Sequence 23, Appli
28	1494.5	50.5	892	5	US-10-982-512-23	Sequence 27, Appli
29	1494.5	50.5	892	5	US-10-982-512-27	Sequence 27, Appli
30	1494.5	50.5	892	5	US-10-433-757-12	Sequence 12, Appli
31	1494.5	50.5	969	4	US-10-415-122-2	Sequence 2, Appli
32	1488.5	50.3	869	4	US-10-415-122-4	Sequence 4, Appli
33	1486.5	50.3	863	4	US-10-072-012-224	Sequence 224, App
34	1486.5	50.3	863	4	US-10-072-012-226	Sequence 226, App
35	1445	48.9	879	3	US-09-976-674-33	Sequence 33, Appli
36	1445	48.9	879	3	US-09-976-674-35	Sequence 35, Appli
37	1445	48.9	879	3	US-10-982-512-33	Sequence 33, Appli
38	1445	48.9	879	5	US-10-982-512-35	Sequence 35, Appli
39	1377.5	46.6	310	3	US-09-993-959-4	Sequence 4, Appli
40	1377.5	46.6	310	4	US-10-825-632-3	Sequence 3, Appli
41	1298.5	43.9	497	3	US-09-833-245-1594	Sequence 1594, Ap
42	1294.5	43.8	497	3	US-09-833-245-1593	Sequence 1593, Ap
43	1187.5	40.2	508	4	US-10-237-271-3	Sequence 3, Appli
44	1187.5	40.2	508	4	US-10-072-012-623	Sequence 623, App
45	1094.5	37.0	832	3	US-09-976-674-29	Sequence 29, Appli

#### ALIGNMENTS

RESULT 1

US-10-825-632-5

; Sequence 5, Application US/10825632

; Publication No. US20040191826A1

; GENERAL INFORMATION:

; APPLICANT: ABBOTT, Catherine Anne

; APPLICANT: GORRELL, Mark Douglas

; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES

; FILE REFERENCE: FCSB-100-Div. 1

; CURRENT APPLICATION NUMBER: US/10/825,632

; CURRENT FILING DATE: 2004-04-15

; PRIOR APPLICATION NUMBER: US 10/070,464

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: PCT/AU00/01085

; PRIOR FILING DATE: 2000-09-11

; PRIOR APPLICATION NUMBER: AU PQ5709

; PRIOR FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: AU PQ2762

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 465

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-825-632-5

Alignment Scores:

Pred. No.: 1.01e-245

Score: 2482.00

Percent Similarity: 100.0%

Best Local Similarity: 100.0%

Query Match: 83.9%

DB:

Length: 465  
Matches: 465  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-825-632-6 (1-1669) x US-10-825-632-5 (1-465)

QY 2 ACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATAATGATTGATGCTGAA 61  
Db 1 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluMetIleAspAlaGlu 20  
QY 62 GGAAGGATCATAGATGTCTAGATAGGAACCTAAATCAACCTTTTGAGATCTTATTTGAA 121  
Db 21 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 40  
QY 122 GGAGTTGAATATATGCCAGAGCTGGATGGACTCTGAGGAAATATATGCTTGGTCCATC 181  
Db 41 GlyValGluTyrIleAlaArgAlaGlyThrThrProGluGlyLysIleAlaTrpSerIle 60  
QY 182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACTGATATATTATC 241  
Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 80  
QY 242 CCGTAGAGATGATGTTATGAAAGGAGAGACTCATGTAGTCAGTCCGCTGATCTGTG 301  
Db 81 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100  
QY 302 ACCGCACTAATTATCATGAAGAAACAACAGACATCTGGATATAATATCCATGACATCTTT 361  
Db 101 ThrProLeuIleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePhe 120  
QY 362 CATGTTTTCCCAAGTCACAGAGGAAATTCAGTTTATTTTGGCTCTGAAATCAAA 421  
Db 121 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 140  
QY 422 ACAGGTTTCCTCATTTATACAAAATACATCTATTTTAAAGGAAAGCAATATAACGA 481  
Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 160  
QY 482 TCCAGTGGTGGCTGCTGCTCCAGTGTATTCAGTGTCTCAAGTGTCTCAAGAGGAGATGCA 541  
Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 180  
QY 542 ATTACAGTGGTGAATGGAGTCTTGGCCGCGATCGATCAATATATCCAGTTGATGAA 601  
Db 181 IleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 200  
QY 602 GTCAGAGGCTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAC 661  
Db 201 ValArgGluValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 220  
QY 662 GTAGTCAGTTACGTAAATCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCAT 721  
Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240  
QY 722 TCTTGTGCTGATCAGTCAGCATCTGACTCTCTTTATAGTAAGTATATAGTAACCAAGAGAT 781  
Db 241 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 260  
QY 782 CCACACTGTGTGCTCTTACCAAGCTATCAAGTCTCAAGATGACCCCACTTGGCAAAACA 841  
Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 280  
QY 842 AAGGAATTTTGGCCCACTTTTGGATTCAGACAGTCTCTTCTGACTATATCTCTCCA 901  
Db 281 LysGluPheThrPalaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300  
QY 902 GAAATTTTCTCTTTTGAAGTACTCTGGATTTACATTTGATGGGATGCTCTACAGCCT 961  
Db 301 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320  
QY 962 CATGATCTACAGCTGGAAAGAAATATCTTACTGTCTGTTCATATATGGTGGCTCCAG 1021  
Db 321 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 340  
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGATCTTCTATCATACAGATACAGGAA 1081  
Db 341 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheIleAspThrGlyTyrThrGlu 360

QY 1082 CGTTATATGGTGCACCTCTGACCGAATGAACAGGGCTATTACTTAGGATCTGTGGCCATG 1141  
Db 361 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 380  
QY 1142 CAACGAGAAAGTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTCTCTGGAT 1201  
Db 381 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 400  
QY 1202 GAGAAATCCATTTTGCACATACCACTATATATTACTAGTGTGAGTTTATAGTGAGGCTGGAAG 1261  
Db 401 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 420  
QY 1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACACAGATAAAGAGTTCCCTGAATCGGGA 1321  
Db 421 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 440  
QY 1322 GAACATTATGAATGATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCACGATTGCT 1381  
Db 441 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 460  
QY 1382 GCTCTAAAGTGATA 1396  
Db 461 AlaLeuLysValIle 465

RESULT 2  
US-09-976-674-1  
; Sequence 1, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Jean-Louis  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-1

Alignment Scores:  
Pred. No.: 1,976-239 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: 3 Gaps: 1

US-10-825-632-6 (1-1669) x US-09-976-674-1 (1-882)

QY 2 ACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATAATGATTGATGCTGAA 61  
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluMetIleAspAlaGlu 337  
QY 62 GGAAGGATCATAGATGTCTAGATAGGAACCTAAATCAACCTTTTGAGATCTTATTTGAA 121  
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357  
QY 122 GGAGTTGAATATATTTGCCAGAGCTGGATGGACTCTGAGGAAATATATGCTTGGTCCATC 181  
Db 358 GlyValGluTyrIleAlaArgAlaGlyThrThrProGluGlyLysIleAlaTrpSerIle 377  
QY 182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTTCACTGAAATATTATTC 241  
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397

QY	242	CGAGTAGAAGATGATGTTATGAAAGGACAGAGACTCATTGAGTCAGTCCCTGATTCTGTG	301	QY	1022	GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAA	1081
Db	398	ProValGluAspValMetGluArgGluArgGluLeuGluSerValProAspSerVal	417	Db	758	ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu	777
QY	302	ACGCCACTAATTATCTATGAAGAAACAAACAGACATCTCTGATAAATATCCATGACATCTTT	361	QY	1082	CGTTATATGGTCACCTCCAGCAGAAATGAACAGGGCTATTACTTAGAGATCTGTGGCCATG	1141
Db	418	ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe	437	Db	778	ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet	797
QY	362	CATGTTTTTCCCAAGTCAAGAGGAAATAGATTTATTTTGGCTCTCTGAAATGCAAA	421	QY	1142	CAGCAGAAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTCGAT	1201
Db	438	HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys	457	Db	798	GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAsp	817
QY	422	ACAGTTTCCTCATTTATACAAATACATCTATTTTAAAGGAAACCAATATAACGA	481	QY	1202	GAGAATGTCATTTTGCACATACACAGTATATTACTAGTATTTTGTAGTGAGGCTGGAAG	1261
Db	458	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	477	Db	818	GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys	837
QY	482	TCCAGTGGTGGGCTGCTCTCCAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCA	541	QY	1262	CCATATGATTACAGATCTATCTCTCAGGAGACACACAGATAAGAGTTCTCGAATCCGGA	1321
Db	478	SerSerGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla	497	Db	838	ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly	857
QY	542	ATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGCATGATCTAATATCCAAAGTTGATGA	601	QY	1322	GAACATTATGAACCTGCATCTTTTGCACACTACCTTCAAGAAAACTTTGGATCAGCTATTGCT	1381
Db	498	IleThrSerGlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGlu	517	Db	858	GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla	877
QY	602	GTGAGAGGCTGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAC	661	QY	1382	GCTCTAAAAGTGATA	1396
Db	518	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr	537	Db	878	AlaLeuLysValIle	882
QY	662	GTAGTCAGTTACGTAAATCCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTACAT	721	RESULT 3			
Db	538	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	557	US-10-054-776-2			
QY	722	TCTTGTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCGAAGAT	781	; Sequence 2, Application US/10054776			
Db	558	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn	577	; Publication No. US20030165818A1			
QY	782	CCACACTGTGTGCTTTCACAGCTATCAAGTCTGAGATGACCCCACTGCGCAACCAACA	841	; GENERAL INFORMATION:			
Db	578	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	597	; APPLICANT: Mark Robert Edbrooke			
QY	842	AAGGATTTTGGCCACCATTTTGGATTTCAGAGTCTCTCTCTGACTATACTCTCTCCA	901	; APPLICANT: Alan Peter Lewis			
Db	598	LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	617	; TITLE OF INVENTION: NOVEL PROTEIN			
QY	902	GAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGGATGCTCTCAAGCCT	961	; FILE REFERENCE: OG1042US			
Db	618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	937	; CURRENT APPLICATION NUMBER: US/10/054,776			
QY	962	CATGATCTACAGCTCGAAGAAATATCTACTGTGTGTTTCATATATGTGTGCTCTCAG	1021	; CURRENT FILING DATE: 2002-01-23			
Db	638	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	657	; NUMBER OF SEQ ID NOS: 2			
QY	1021	-----	1021	; SOFTWARE: PatentIn version 3.0			
Db	658	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	677	; SEQ ID NO 2			
QY	1021	-----	1021	; TYPE: PRT			
Db	678	SerLeuGlyTyrValValValIleAspAsnArgLysCysHisArgLysLeuLys	697	; ORGANISM: Homo sapiens			
QY	1021	-----	1021	US-10-054-776-2			
Db	698	PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly	717	Alignment Scores:			
QY	1021	-----	1021	Pred. NO.:			
Db	718	LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly	737	Length: 882			
QY	1021	-----	1021	Matches: 882			
Db	738	TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	757	Mismatches: 0			
				Conservative: 0			
				Best Local Similarity: 82.3%			
				Query Match: 81.9%			
				Indels: 100			
				Gaps: 1			
				DB:			
				US-10-825-632-6 (1-1669) x US-10-054-776-2 (1-882)			
QY	2	ACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATGTGAGAAATAATGATTGATCGTAA	61	QY	318	ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu	337
Db				Db	62	GGAAGATCATAGATCTCATAGTAAGGAACCTAACTTTCACCTTTTGAGATCTCTATTGAA	121
QY				Db	338	GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu	357
QY				QY	122	GGAAGTTGAATATATTCGCCAGAGCTGGATGCACTCTCTGAGGGAATAATGCTTGTCTCATC	181
Db				Db	358	GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle	377
QY				QY	182	CTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAAATATTATTC	241
Db				Db	378	LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle	397

QY 242 CCAAGTAAAGATGATGTTATGTAAGAAAGGCAGAGACTCATTGAGTCAGTGGCTGATCTGTG 301  
Db 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 AGCCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361  
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437  
QY 362 CATGTTTTTCCCAAGTCAACAAGAGAAATTCAGTTTATTTTTCCTCTCGAATCAAA 421  
Db 438 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGGTTTCCTCATTTATACAAATACATCTATTTTAAAGAAAGCAAAATATAACGA 481  
Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysIleAsnG 477  
QY 482 TCCAGTGGTGGCTGCCTGCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA 541  
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497  
QY 542 ATTACAGTGGTGAATGGAGTCTTGGCCGGCATGGATCTAATATCCAGTTGATGAA 601  
Db 498 IleThrSerGlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTCAAGAGCTGGTATATTTGAGGCAACCAAGACTCCCTTTAGAGCATCCTGTRAC 661  
Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyr 537  
QY 662 GTAGTCAGTTACGTAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACAT 721  
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
QY 722 TCTGTCTGATCAGTCAGCACTGTGACTCTTTTATAAGTAAAGTATATAGTAACAGAGAAAT 781  
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
QY 782 CCACACTGTGTCTCCTTTACAAAGCTATCAAGTCTCTGAAGTACCCCACTTGCAGAAACA 841  
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
QY 842 AAGGAATTTGGCCACCATTTTGATTCAGCAGCTCTCTCTGACTATATCTCTCCA 901  
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
QY 902 GAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCT 961  
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
QY 962 CATGATCTACGCTGGAAGAAATATCTACTCTGCTGTTTCATATATGTTGCTCTCAG 1021  
Db 638 HisAspGluProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657  
QY 1021 ----- 1021  
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677  
QY 1021 ----- 1021  
Db 678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
QY 1021 ----- 1021  
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
QY 1021 ----- 1021  
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021  
Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1081

Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777  
QY 1082 CGTATATGGGTACCCCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCATG 1141  
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797  
QY 1142 CAAGCAGAAAAAGTTCCCTCTCTGAACCAAAATCGTTTACTGCTTACATGGTTCCTGGAT 1201  
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817  
QY 1202 GAGAATGTCATTTTGCACATACCAAGTATATTACTAGTTTTTTAGTGAGGCTGGAAG 1261  
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
QY 1262 CCATATGATTACAGATCTATCTCCAGAGAGACACAGCATAAGAGTTCTCGAATCGGA 1321  
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857  
QY 1322 GACATTTATGAAGTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGTATTGCT 1381  
Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877  
QY 1382 GCTCTAAAAGTGATA 1396  
Db 878 AlaLeuLysValIle 882  
RESULT 4  
US-10-170-789-38  
; Sequence 38, Application US/10170789  
; Publication No. US20030180930A1  
; GENERAL INFORMATION:  
; APPLICANT: Rachel E. Meyers  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Weich, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,  
; FILE REFERENCE: 10448-191001  
; CURRENT APPLICATION NUMBER: US/10/170,789  
; CURRENT FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/882,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/212,078  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/934,406  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US01/26052  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,740  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 09/861,801  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: PCT/US01/16549  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 60/205,508  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 09/801,267  
; PRIOR FILING DATE: 2001-03-06  
; PRIOR APPLICATION NUMBER: PCT/US01/07138  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: US 60/187,454  
; PRIOR FILING DATE: 2000-03-07

;; PRIOR APPLICATION NUMBER: US 09/829,671  
;; PRIOR FILING DATE: 2001-04-10  
;; PRIOR APPLICATION NUMBER: PCT/US01/40483  
;; PRIOR FILING DATE: 2001-04-11  
;; PRIOR APPLICATION NUMBER: US 60/197,508  
;; PRIOR FILING DATE: 2000-04-18  
;; PRIOR APPLICATION NUMBER: US 09/961,721  
;; PRIOR FILING DATE: 2001-09-24  
;; PRIOR APPLICATION NUMBER: PCT/US01/29904  
;; PRIOR FILING DATE: 2001-09-24  
;; PRIOR APPLICATION NUMBER: US 60/235,023  
;; PRIOR FILING DATE: 2000-09-25  
;; PRIOR APPLICATION NUMBER: US 10/045,367  
;; PRIOR FILING DATE: 2001-11-07  
;; PRIOR APPLICATION NUMBER: US 60/246,561  
;; PRIOR FILING DATE: 2000-11-07  
;; PRIOR APPLICATION NUMBER: US 09/801,275  
;; PRIOR FILING DATE: 2001-03-06  
;; PRIOR APPLICATION NUMBER: PCT/US01/07074  
;; PRIOR FILING DATE: 2001-03-05  
;; PRIOR APPLICATION NUMBER: US 60/187,420  
;; PRIOR FILING DATE: 2000-03-07  
;; NUMBER OF SEQ ID NOS: 63  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 38  
;; LENGTH: 882  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-170-789-38

## Alignment Scores:

Pred. No.: 1 97e-239 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: Gaps: 1

US-10-825-632-6 (1-1669) x US-10-170-789-38 (1-882)

QY 2 ACAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGAA 61  
DB 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337  
QY 62 GGAAGGATCATAGATCATAGATAAGGAACATAATCAACCTTTTGAGATTCTATTGAA 121  
DB 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGlu 357  
QY 122 GAGTTGAATATATATCCAGAGCTGGATGAGTCTCTGAGGGAATAATGCTTGTCCATC 181  
DB 358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377  
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATATTATTC 241  
DB 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
QY 242 CCAGTACAGATGATGTTATGGAAGCAGAGACTCATTGAGTCAGTGCCTGATTCGTG 301  
DB 398 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACGCCATAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361  
DB 418 ThrProLeuIleIleTyrGluIleThrThrAspIleTyrIleAsnIleHisAspIlePhe 437  
QY 362 CATGTTTTCCTCCCAAGTACGGAAGGAAATGAGTTATTTTTCCTCTGAAATGCAAA 421  
DB 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGTTTCGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAAAATATAACGA 481  
DB 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 477  
QY 482 TCCAGTGTGGCTCGCTCTCCCAAGTGATTTCAAGTGCTTATCAAGAGGAGATAGCA 541

DB 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAla 497  
QY 542 ATTACCAAGTGGTCAATGGGAAGTCTTGTGCGGATGATCTAATATATCAAGTTGATGAA 601  
DB 498 IleThrSerGlyLutrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTCAGAAGCTGGTATATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAC 661  
DB 518 ValArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyr 537  
QY 662 GTAGTCAGTTTACGTAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGTACTCATCAT 721  
DB 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
QY 722 TCTTGTCGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781  
DB 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
QY 782 CCACACTGTGTGCTCTTTCAAGCTATCAAGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCA 841  
DB 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
QY 842 AAGGAATTTTGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTCTGACTATCTCTCTCA 901  
DB 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
QY 902 GAAATTTTCTTTTGAAGTACTACTGATTTTACATTTGATGATGATGATGATGATGATGAT 961  
DB 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
QY 962 CATGATCTACAGCTCGAAGAAATATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021  
DB 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657  
QY 1021 ----- 1021  
DB 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677  
QY 1021 ----- 1021  
DB 678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
QY 1021 ----- 1021  
DB 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
QY 1021 ----- 1021  
DB 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021  
DB 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTCTATTCTGGGGCCCAAGTCACTCTGTGATCTCTATGATACAGGATACACGAA 1081  
DB 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777  
QY 1082 CGTTATATGGTTCACCTCGACCAAGTGAAGGCTGATTTACTAGGATCTGTGGCCATG 1141  
DB 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797  
QY 1142 CAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTCTCTTACTCTTACTCTCTCTCTCTCT 1201  
DB 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817  
QY 1202 GAGAAATGTCATTTTGCATACATACATATATCTAGTCTTTTGTAGTGGGCTGGAAG 1261  
DB 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
QY 1262 CCATATGATTTACAGATCTATCTCAGGAGACAGCATAGAGTCTCTCAATCGGGA 1321



```
Db      838 ProTyrAspLeuGluInileTyrProGlnGluArgHisSerIleArgValProGluSerGly 857
Qy      1322 GAACATTATGAACCTGATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGCTATTGCT 1381
Db      858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAla 877
Qy      1382 GCTCTAAAGTGATA 1396
Db      878 AleLeuLysValIle 882

RESULT 5
US-10-311-035-9
; Sequence 9 Application US/10311035
; Publication No US20040023243A1
; GENERAL INFORMATION: INCYTE GENOMICS, INC.
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Deepoojiya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: proteases
; FILE REFERENCE: PI-012356-CT
; CURRENT APPLICATION NUMBER: US/10/311,035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1
US-10-311-035-9

Alignment Scores:
Pred. No.: 1,97e-239 Length: 882
Score: 2422.00 Matches: 465
Percent Similarity: 82.3% Conservative: 0
Best Local Similarity: 82.3% Mismatches: 0
Query Match: 81.9% Indels: 100
DB: 4 Gaps: 1

US-10-825-632-6 (1-1669) x US-10-311-035-9 (1-882)
Qy      2 ACAGGTACAGCAATCCCTAAAGTCACCTTTTAAGATGTTCAGAAATATGATTCCTGAA 61
Db      318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337
```

```
Qy      62 GGAAGGATCATAGATGTCTATAGATAAGAACTAAATCAACCTTTTGAGATTCATTGAA 121
Db      338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
Qy      122 GGATTGAATATATGTCAGAGCTGGATGGATCCCTGAGGGAAAAATATGCTTGGTCCATC 181
Db      358 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 377
Qy      182 CTACTAGATCGCTCCAGAGCTCGCCCTACAGATAGTGTGATCTCTACCTGAAATATATTATC 241
Db      378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
Qy      242 CCAGTAGAAGATGATGTTATGGAAGGCAGAGACTCATTGAGTCAGTCCCTGATTCGTGTG 301
Db      398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
Qy      302 ACGCCACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTT 361
Db      418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrrPleAenIleHisAspIlePhe 437
Qy      362 CATGTTTTTCCCAAAAGTCACGAAGAGAAATTTGAGTTTATTTTTTGGCTCTGAAATGAAA 421
Db      438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457
Qy      422 ACAGGTTTCGTCATTTATACAAAATTACATCTATTTTAAAGGAAGCAAAATATAACGA 481
Db      458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477
Qy      482 TCCAGTGGTGGGCTGCTCCCAAGTATTCAGTGTCTCTATCAAGAGGAGATAGCA 541
Db      478 SerSerGlyLeuProLysProSerAspPheLysCysProIleLysGluGluIleAla 497
Qy      542 ATTACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCAAGTTGATGAA 601
Db      498 IleThrSerGlyLeuTrpGluValLeuGlyArgHisGlySerAenIleGlnValAspGlu 517
Qy      602 GTCAGAAGCTGGTATATTTTGAAGGCCCAAGAACTCCCTTTAGAGCATCACCTGTAC 661
Db      518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537
Qy      662 GTAGTCAGTTACGTAAATCCTGGAGAGGTGAACAAGGTGACTGACCGTGGCTACTCATCAT 721
Db      538 ValValSerTyrValAsnProGlyValThrArgLeuThrAspArgGlyTyrSerHis 557
Qy      722 TCTTGCTGCATCAGTCAGCATGTGACTTCTTTATAAGTAAGTATAGTAAACCAAGAAAT 781
Db      558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577
Qy      782 CCACACTGTGTCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACA 841
Db      578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
Qy      842 AAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTCTGACTATATCTCTCCA 901
Db      598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
Qy      902 GAAATTTTCTCTTTTGAAGTACTACTCGATTTTACATTTGATGGATGGCTCTCAAGCCT 961
Db      618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
Qy      962 CATGATCTACAGCTGGAAAGAAATACCTACTGCTGTGTCTATATATGCTGCTCTCAG 1021
Db      638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657
Qy      1021 ----- 1021
Db      658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAla 677
Qy      1021 ----- 1021
Db      678 SerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697
Qy      1021 ----- 1021
```

Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleuLeuAspGlnValGluGly 717  
QY 1021 ----- 1021  
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021  
Db 738 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTGTATTGTGGGGCCCCAGTCCTCTGTGTGATCTTCTATCATACAGGATACACGGAA 1081  
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777  
QY 1082 CTTTATATGGTCCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCCATG 1141  
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797  
QY 1142 CAAGCAGAAAGTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTGGAT 1201  
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817  
QY 1202 GGAATGTCCTATTTGCACATACCATATATTACTAGTCTTTTGTAGTGGGGCTGGAAG 1261  
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
QY 1262 CCATATGATTTACAGATCTCTCTCAGGAGAGACAGCATTAAGAGTTCTCGAATCGGGA 1321  
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857  
QY 1322 GAACATATATGAAGTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGTATTGCT 1381  
Db 858 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877  
QY 1382 GCTCTAAAGTGATA 1396  
Db 878 AlaLeuLysValIle 882  
RESULT 6  
US-10-072-012-622  
; Sequence 622, Application US/10072012  
; Publication No. US20040033493A1  
; GENERAL INFORMATION:  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shinkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Anderson, David W.  
; APPLICANT: Hastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier Jr, Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Pena, Carol E. A  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Drosse, William M.  
; APPLICANT: Alsbrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-258  
; CURRENT APPLICATION NUMBER: US/10/072,012  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/265,102  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: 60/265,514  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,412  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,395  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/266,406  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/266,767  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: 60/267,057  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/266,975  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/267,459  
; PRIOR FILING DATE: 2001-02-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1391  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 622  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-072-012-622

Alignment Scores:  
Pred. No.: 1.97e-239 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: 4 Gaps: 1

US-10-825-632-6 (1-1669) x US-10-072-012-622 (1-882)

QY 2 ACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAGATCTCAGAAATAATGATTGATCTGAA 61  
Db ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337  
QY 62 GGAAGGATCATAGATGTCTAGATAGAACTAAATCAACCTTTTTCAGATTCTATTGAA 121  
Db GlyArgIleLeuAspValIleAspLysGluLeuIleGlnProPheGluLeuLeuPheGlu 357  
QY 122 GGAGTTGAATATATTCAGAGCTGGATGGATCTCTGAGGAAAATATGCTTGGTCCATC 181  
Db GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle 377  
QY 182 CTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTACCTGGAATTATTATTC 241  
Db LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
QY 242 CCAGTAGAAGATGATGTATTGGAAGGACAGAGACTCATTGACTGCTGCTGATCTCTG 301  
Db ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACGCCACTAATTATCTATGAAGAAACAAGACATCTGATATAATATCATGACATCTTT 361  
Db ThrProLeuIleIleTyrGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437  
QY 362 CATGTTTTTCCCAAGTCACGAAGAGAAATGAGTTTATTATTTTTCCTCTGAAATGCAA 421  
Db HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGTTTCCGTCAATTTATACAAATTAATCATCTATTATTTAAAGGAAAGCAAAATATA 481  
Db ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
QY 482 TCCAGTGTGGCTGCTGCTCCAGTGATTTCAAGTGATTTCAAGTCTCATCAAGAGAGATACA 541  
Db SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluLeuIleAla 497



Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTGAGAGGCTGGTATATTTTGAAGCGACCAAGACTCCCTTTAGAGCATCACCTGTAC 661  
Db 518 ValArgGluValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 537  
QY 662 GTAGTCAGTTACCTAAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACAT 721  
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
QY 722 TCTGTGTCGATCAGTCAGCACCTGACTCTCTTTATAGTAAGTATAGTAAACAGAGAT 781  
Db 558 SerCysHisSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
QY 782 CCACACTGTGTGTCCTTTACAGCTATCAAGCTCTCAAGATGACCCCACTTGCAAAACA 841  
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
QY 842 AAGGAATTTGGGCCACCAATTTGGATTTCAGAGTCTCTCTCTGACTATCTCTCTCCA 901  
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
QY 902 GAATTTCTCTTTCAAGTACTACTGGATTACATTTGATGGATGCTCTCAGACCT 961  
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
QY 962 CATGATCTACAGCTCGAAAGAAATATCTACTGTCTGTTTCATATATGTGGTCTCAG 1021  
Db 638 HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGlyProGln 657  
QY 1021 ----- 1021  
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 677  
QY 1021 ----- 1021  
Db 678 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
QY 1021 ----- 1021  
Db 698 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
QY 1021 ----- 1021  
Db 718 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021  
Db 738 TrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTGCTATTCTGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGAA 1081  
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 777  
QY 1082 CGTTATATGGTCACTCACCAGATGAACAGGCTATTACTTAGGATCTGTGGCCATG 1141  
Db 778 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 797  
QY 1142 CAACGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTCTCGAT 1201  
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817  
QY 1202 GAGAACTCTCAATTTGCACATACAGTATATTACTAGTATTTTGTAGTGGCTGGAAG 1261  
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837  
QY 1262 CCAATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCTTGAATCGGA 1321  
Db 838 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 857  
QY 1322 GAACATTTGAAGTCACTCTTTGCACTACCTTCAAGAAACCTTGGATCAGTATTGCT 1381  
Db 858 GluHisTyrGluLeuHisLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 877

QY 1382 GCTCTAAAAAGTGATA 1396  
Db 878 AlaLeuLysValIle 882  
RESULT 8  
US-10-825-632-1  
; Sequence 1, Application US/10825632  
; Publication No. US20040191826A1  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: CORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: FCSB-100-Div. 1  
; CURRENT APPLICATION NUMBER: US/10/825,632  
; CURRENT FILING DATE: 2004-04-15  
; PRIOR APPLICATION NUMBER: US 10/070,464  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-825-632-1

Alignment Scores:  
Pred. No.: 1,97e-239 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: 4 Gaps: 1

US-10-825-632-6 (1-1669) x US-10-825-632-1 (1-882)

QY 2 ACAGGTACAGCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGATGCTGAA 61  
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337  
QY 62 GGAGGATCATAGTCTCATAGTAGAGAACTAAATCAACCTTTTGGATTTCTATTGAA 121  
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357  
QY 122 GGAGTTGAATATATTGCCAGAGCTGGATGCACTCCTCGAGGGAATAATATGCTTGGTCCATC 181  
Db 358 GlyValGluTyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIle 377  
QY 182 CTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGTGATCTCACCTGAATTTATTATC 241  
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
QY 242 CCAGTAGAAGATCATGTTATGGAAAGCAGACACTCATTGAGTCAGTCGCTGATTCTGTG 301  
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACGCCACTAATATTCTATGAAGAAACAAGACATCTCGATAAATATATCCATGACATCTTT 361  
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437  
QY 362 CATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTTATTTTGGCTCTGAAATCGAA 421  
Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457  
QY 422 ACAGTTTCTGTCATTTATACAAATATACATCTATTATTAAGAAAGCAAAATATATAACGA 481  
Db 458 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 477

QY 482 TCACGTGGTGGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCA 541  
Db 478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluLeuAla 497  
QY 542 ATTACCAAGTGGTGAAGTCTTGGCCGCGCATGGATCTAATATCCAAAGTGTAGAA 601  
Db 498 IleThrSerGlyGluTrpGluValLeuGlyAsgHisGlySerAsnIleGlnValAspGlu 517  
QY 602 GTACAGAGGCTGATATTTTGAAGCACCACCAAGACTCCCTTTAGAGCATCCTGTAC 661  
Db 518 ValArgGluValThrPheGluGlyThrLysAspSerProLeuGluHisLysLeuThr 537  
QY 662 GTAGTCAAGTAACTCTGGAGAGGTGAACAAGCTGACTGACCGTGGCTACTACAT 721  
Db 538 ValValSerThrValAsnProGlyGluValThrArgLeuThrAspArgGlyThrSerHis 557  
QY 722 TCTTGCTGCATAGTCAGCATGCTGACTCTTTTATAAGTAAAGTAACTGAACGAAGAT 781  
Db 558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysThrSerAsnGlnLysAsn 577  
QY 782 CCACACTGCTGCTCCCTTTACAAGCTATCAAGTCTCAAGATGACCCCAACTTGCAAAACA 841  
Db 578 ProHisCysValSerLeuThrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
QY 842 AAGCAATTTTGGCCCACTTTTGGATTTCAGAGTCTCTCTCTGACTATCTCTCCA 901  
Db 598 LysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuProAspThrProPro 617  
QY 902 GAAATTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGCCT 961  
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuThrGlyMetLeuThrLysPro 637  
QY 962 CATGATCTACAGCTCGGAAGAAATATCTCTGCTGCTGCTCATATATGCTGCTCTCAG 1021  
Db 638 HisAspLeuGlnProGlyLysLysThrProThrValLeuPheIleThrGlyGlyProGln 657  
QY 1021 ----- 1021  
Db 658 ValGlnLeuValAsnAsnArgPheLysGlyValLysThrPheArgLeuAsnThrLeuAla 677  
QY 1021 ----- 1021  
Db 678 SerLeuGlyThrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 697  
QY 1021 ----- 1021  
Db 698 PheGluGlyAlaPheLysThrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 717  
QY 1021 ----- 1021  
Db 718 LeuGlnThrLeuAlaSerArgThrAspPheIleAspLeuAspArgValGlyIleHisGly 737  
QY 1021 ----- 1021  
Db 738 TrpSerThrGlyGlyThrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 757  
QY 1022 GTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAA 1081  
Db 758 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheThrLysAspThrGlyThrGlu 777  
QY 1082 CGTTATATGGTCAACCTGACAGAAATGAACAGGCTATTACTTAGATCTGCGCCATG 1141  
Db 778 ArgThrMetGlyHisProAspGlnAsnGluGlnGlyThrLeuGlySerValAlaMet 797  
QY 1142 CAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCTACATGCTTCCTGGAT 1201  
Db 798 GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 817  
QY 1202 GAGATGCTCAATTTGCAATACAGATATATCTAGATTTTGTAGTGGCTGGAAG 1261  
Db 818 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 837

QY 1262 CCATATGATTTACAGATCTATCTCTCAGGAGACACACACATTAAGAGTTCTCTGAATCGGA 1321  
Db 838 ProThrAspLeuGlnIleThrProGlnGluArgHisSerIleArgValProGluSerGly 857  
QY 1322 GAACTATTAAGTCACTCTTTTGCCTACTACCTTTCAAGAAAACCTTGGATCACGATTTGCT 1381  
Db 858 GluHisThrGluLeuHisLeuLeuHisLysThrLeuGlnGluAsnLeuGlySerArgIleAla 877  
QY 1382 GCTCTAAAAGTGATA 1396  
Db 878 AlaLeuLysValIle 882  
RESULT 9  
US-10-982-512-1  
; Sequence 1, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; PRIOR FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-982-512-1  
Alignment Scores:  
Pred. No.: 1,97e-239 Length: 882  
Score: 2422.00 Matches: 465  
Percent Similarity: 82.3% Conservative: 0  
Best Local Similarity: 82.3% Mismatches: 0  
Query Match: 81.9% Indels: 100  
DB: 5 Gaps: 1  
US-10-825-632-6 (1-1669) x US-10-982-512-1 (1-882)  
QY 2 ACAGGTACAGCAAACTCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTGCTGAA 61  
Db 318 ThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu 337  
QY 62 GGAAGGATCATAGATGTCATAGATAAGGAACTAATTAACCTTTTGAGATTTCTATTGAA 121  
Db 338 GlyArgIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357  
QY 122 GGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGTCCTC 181  
Db 358 GlyValGluThrIleAlaArgAlaGlyThrThrProGluGlyLysThrAlaTrpSerIle 377  
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCCACCTGAATATTATTC 241  
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397  
QY 242 CCAGTAGAGATCATGTTTATGGAAAGGCGAGACTCATTGATGCTGCTGCTGCTG 301  
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417  
QY 302 ACGCCCACTAATTTCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTT 361  
Db 418 ThrProLeuIleGlyThrGluThrThrAspIleThrIleAsnIleHisAspIlePhe 437  
QY 362 CATGTTTTCCTCCCAAGTCACAGAGGAAATTCAGTTTATTTTGTGCTCTGATGCAAA 421

Db	438	HisValPheProGlnSerHisGluGluGluLeuPheAlaSerGluCysLys	457	QY	1202	GAGATGTCATTTGCACATACCAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAG	1261
QY	422	ACAGGTTTCCGTCAATTTATATAAAATTTACATCTATTATTAAGGAAAGCAAAATATATAACGA	481	Db	818	GluAsnValHisPheAlaHisThrSerIleuLeuLeuSerPheLeuValArgAlaGlyLys	837
Db	458	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	477	QY	1262	CAATATGATTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGATTCCTGAATCGGA	1321
QY	482	TCCAGTGTGGCTGCTGCTCCCAAGTGATTTCAAGTGCCTTATCAAGAGAGATACGA	541	Db	838	ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly	857
Db	478	SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluLeuIleAla	497	QY	1322	GAACATTATGAATGCGATCTTTTGCATCTACCTTCAGAAAAACCTTCGATCAGTATTGCT	1381
QY	542	ATTACAGTGTGAATGGAGTCTTGGCGGCATGGATCTAAATATCAAGTCTCATCAA	601	Db	858	GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla	877
Db	498	IleThrSerGlyGluIleProGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu	517	QY	1382	GCTCTAAAAGTCATA	1396
QY	602	GTCAGAAGCTGGTATATTTTGAAGGCCAACCAAGATCTCCCTTTAGAGCATCACCTGTAC	661	Db	878	AlaLeuLysValIle	882
Db	518	ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLysTyr	537	RESULT 10			
QY	662	GTAGTCAGTACGTAAATCTCTGGAGAGGTGACAAAGCTGACCGTGGCTACTCAT	721	US-10-072-012-621			
Db	538	ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis	557	; Sequence 621, Application US/10072012			
QY	722	TCTTGCTGCATCAGTCAGCATGTGACTTCTTTATAGTAAAGTATAGTAAACAGAGAT	781	; Publication NO. US20040033493A1			
Db	558	SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn	577	; GENERAL INFORMATION:			
QY	782	CGACTGTGTGCTCCTTTACAGCTATCAAGTCTCAAGTCAAGCACCACCACTTGCACAAACA	841	; APPLICANT: Tchernev, Velizar			
Db	578	ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr	597	; APPLICANT: Spytek, Kimberly			
QY	842	RAGGAATTTGGGCCACCATTTTGATTCAGAGCTCTCTCTGACTATATCTCTCTCA	901	; APPLICANT: Zerhusen, Bryan			
Db	598	LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro	617	; APPLICANT: Patturajan, Meera			
QY	902	GAATTTCTCTTTTGAAGTACTACTGGATTACATTTGATGGGATGCTCTACAGCT	961	; APPLICANT: Shinkets, Richard			
Db	618	GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro	637	; APPLICANT: Li, Li			
QY	962	CATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTCATATATATGTTGCTCTCAG	1021	; APPLICANT: Gangolli, Esha			
Db	638	HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	657	; APPLICANT: Fadigar, Muralidhara			
QY	1021	-----	1021	; APPLICANT: Anderson, David W.			
Db	658	ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla	677	; APPLICANT: Rastelli, Luca			
QY	1021	-----	1021	; APPLICANT: Miller, Charles E.			
Db	678	SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys	697	; APPLICANT: Gerlach, Valerie			
QY	1021	-----	1021	; APPLICANT: Taupier Jr, Raymond J.			
Db	698	PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly	717	; APPLICANT: Gusev, Vladimir Y.			
QY	1021	-----	1021	; APPLICANT: Colman, Steven D.			
Db	718	LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly	737	; APPLICANT: Wolenc, Adam R.			
QY	1021	-----	1021	; APPLICANT: Pena, Carol E. A			
Db	738	TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg	757	; APPLICANT: Furtak, Katarzyna			
QY	1022	GTTCCTATTGCTGGGCCCCAGTCACTCTGGGATCTTCTATGATACAGGATACACGAA	1081	; APPLICANT: Grosse, William M.			
Db	758	ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu	777	; APPLICANT: Alsobrook II, John P.			
QY	1082	CGTTATATGGTCACTCACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATG	1141	; APPLICANT: Lepley, Denise M.			
Db	778	ArgTyrMetGlyHisProAspGlnAsnGlnGlnTyrTyrTyrLeuGlySerValAlaMet	797	; APPLICANT: Rieger, Daniel K.			
QY	1142	CAAGCAGAAAGTTCCCTCTCAACCAAAATCGTTTACTGCTCTTACATGGTTCTCGAT	1201	; APPLICANT: Burgess, Catherine E.			
Db	798	GlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp	817	; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same			
				; FILE REFERENCE: 21402-258			
				; CURRENT APPLICATION NUMBER: US/10/072,012			
				; CURRENT FILING DATE: 2002-01-31			
				; PRIOR APPLICATION NUMBER: 60/265,102			
				; PRIOR FILING DATE: 2001-01-30			
				; PRIOR APPLICATION NUMBER: 60/265,514			
				; PRIOR FILING DATE: 2001-01-31			
				; PRIOR APPLICATION NUMBER: 60/265,517			
				; PRIOR FILING DATE: 2001-01-31			
				; PRIOR APPLICATION NUMBER: 60/265,412			
				; PRIOR FILING DATE: 2001-01-31			
				; PRIOR APPLICATION NUMBER: 60/265,395			
				; PRIOR FILING DATE: 2001-01-31			
				; PRIOR APPLICATION NUMBER: 60/266,406			
				; PRIOR FILING DATE: 2001-02-02			
				; PRIOR APPLICATION NUMBER: 60/266,767			
				; PRIOR FILING DATE: 2001-02-05			
				; PRIOR APPLICATION NUMBER: 60/267,057			
				; PRIOR FILING DATE: 2001-02-07			
				; PRIOR APPLICATION NUMBER: 60/266,975			
				; PRIOR FILING DATE: 2001-02-07			
				; PRIOR APPLICATION NUMBER: 60/267,459			
				; PRIOR FILING DATE: 2001-02-08			
				; Remaining Prior Application data removed - See File Wrapper or PALM.			
				; NUMBER OF SEQ ID NOS: 1391			
				; SOFTWARE: PatentIn Ver. 2.1			
				; SEQ ID NO 621			

```
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Mus musculus
us-10-072-012-621

Alignment Scores:
Pred. No.: 1,49e-229      Length: 883
Score: 2326.00           Matches: 443
Percent Similarity: 80.2% Conservative: 10
Best Local Similarity: 78.4% Mismatches: 12
Query Match: 78.7%      Indels: 100
DB: 4                      Gaps: 1

US-10-825-632-6 (1-1669) x US-10-072-012-621 (1-883)

QY 2 ACAGGTACGCAATCTTAAGTCTCTTTAAGATGTCAGAAATAATGATGATGCTGAA 61
Db 319 ThrGlyThrAlaAsnProValThrPheLysMetSerGluIleValValAspAlaA 338
QY 62 GNAAGGATCATGATGTCATAGATAAGGAACTAATCAACTTTTGAGATTCTATTGAA 121
Db 339 GlyGlyIleLeuAspValIleAspLysGluLeuValGlnProPheGluIleLeuPheGlu 358
QY 122 GGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGGTCCATC 181
Db 359 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysHisAlaTrpSerIle 378
QY 182 CTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCACCTGAATATTATTATC 241
Db 379 LeuLeuAspArgSerGlnThrHisLeuGlnIleValLeuIleSerProGluLeuPheIle 398
QY 242 CCAGTAGAAGATGATGTTATGGAAGAACACAGACATCTGATATATATCATCATCATCTT 301
Db 399 ProValGluAspAlaMetAspArgGlnArgLeuIleGluSerValProAspSerVal 418
QY 302 AGCCACATAATATCTATGAAGAACACAGACATCTGATATATATATATATATATATAT 361
Db 419 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 438
QY 362 CATGTTTTTCCCAAGTCAGAGAGAAATGAGTTATTTTCCCTCTGAATGCAAA 421
Db 439 HisValPheProGlnThrHisGluAspGluIleGluPheIleAspGluSerLysArg 458
QY 422 ACAGGTTTCCGCTATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAACGA 481
Db 459 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysArg 478
QY 482 TCAGTGTGGTGGCTGCTCCAGTGATTTCAAGTGCTCTATCAAGAGAGATAGCA 541
Db 479 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleThr 498
QY 542 ATTACAGTGTGTAATGGAGTTCCTTGGCCGCATGGATCTAATATATCAAGTTGATGAA 601
Db 499 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleTrpValAspGlu 518
QY 602 CTCAGAGGCTGGTATATTTTGAAGGCCAACCAAGACTCCCTTTTAGAGCATCACCTGTAC 661
Db 519 AlaArgLysLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLysTyr 538
QY 662 GTAGTCAGTACGTAAATCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCATCAT 721
Db 539 ValThrSerTyrAlaAsnProGlyGluValValArgLeuThrAspArgGlyTyrSerHis 558
QY 722 TCTTGTGCTACAGTACAGTCTGACTCTCTTTAAGTAAGTATAGTAACACAGAGAAAT 781
Db 559 SerCysCysLeuSerArgHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 578
QY 782 CCACACTGTGTGCTCTTTACAGCTATCAAGTCTGAGATGACCACTTGGCAAAACA 841
Db 579 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProValHisLysThr 598
QY 842 AAGGAATTTGGGCCCACTTTTGGATTTCAGAGGCTCTCTCTCTGACTATATCTCTCCA 901
```

```
RESULT 11
US-10-275-505-2
; Sequence 2, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
```

```
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DELEGANE, Angelo M.; LAL, Preeti G.
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.
; APPLICANT: NGUYEN, Darniel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Junning; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dying Aina M.
```



Db	233	SerSerGlyGlyLeuProAlaProThr	-----C	241
Qy	542	ATTACCACTGGTGAATGGAGTT	-----C	571
Db	242	ValThr	-----T	258
Qy	572	-----CGGCATGGATCTAATATTC	-----C	625
Db	259	CysValThrHisIleValGluIleGlnValAspGluValArgArgLeuValTyPheGlu	-----C	278
Qy	626	GCACCAAAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGTAGTCAGTAACTTAACTTGA	-----C	685
Db	279	GlyThrTyAspSerProLeuGluHisIleLeuTyValValSerTyValAsnProGly	-----C	298
Qy	686	GAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTCTTGTGTCATCAGTCAGTCAGTGT	-----C	745
Db	299	GluValThrArgLeuThrAspArgGlyTySerHisSerCysAlaIleSerGlnHisCys	-----C	318
Qy	746	GACTTCTTTTAAAGTATAGTAAACCAAGAAATCCACACTGTGTGTCTCCTTTTACAA	-----C	808
Db	319	AspPhePheIleSerLysTySerAsnGlnLysAsnProHisCysValSerLeuTyLys	-----C	338
Qy	806	CTATCAAGTCCTCGAAGATGACCCAACTTGCAAACAAAGAAATTTGGGCCACCAATTTG	-----C	865
Db	339	LeuSerSerProGluAspAspProThrCysLysThrLysGluPheThrAlaThrIleLeu	-----C	358
Qy	866	GATTCAAGAGTCCTCTTCTGACTATATCTCTCCAGAAAATTTCTCTTTTGAAGTACT	-----C	925
Db	359	AspSerAlaGlyProLeuProAspTyThrProProGluIlePheSerPheGluSerThr	-----C	378
Qy	926	ACTGGATTACATTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGGAAGAAA	-----C	985
Db	379	ThrGlyPheThrLeuTyGlyMetLeuTyLysProHisAspLeuGlnProGlyLysLys	-----C	398
Qy	986	TATCTCTAGTCGTGTTTCATATATGGTGGTCTTCAG	-----C	1021
Db	399	TyrProThrValLeuPheIleTyGlyGlyProGlnValGlnLeuValAsnAsnArgPhe	-----C	418
Qy	1021	-----C	-----C	1021
Db	419	LysGlyValLysTyPheArgLeuAsnThrLeuAlaSerLeuGlyTyValValValVal	-----C	438
Qy	1021	-----C	-----C	1021
Db	439	IleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyLys	-----C	458
Qy	1022	-----GTTGCTATTGCTGGGGCCAGTCACCTGTGTGATCTTCTATGATACAGGATACAG	-----C	1078
Db	459	MetValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyAspThrGlyTyThr	-----C	478
Qy	1079	GAACGTATATGGGTACCTCCAGACAGATGACAGGCGCTATTACTTAGGATCTGTGCC	-----C	1138
Db	479	GluArgTyMetGlyHisProAspGlnAsnGlnGlnGlyTyTyLeuGlySerValAla	-----C	498
Qy	1139	ATGCAAGCAGAAAAAGTTCCCTCTGACCAAAATCGTTTACTCTCTTACATGTTCTCTG	-----C	1198
Db	499	MetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeu	-----C	518
Qy	1199	GATGAGAATGTCCATTTTGCACATACCATATATTACTGAGTTTTTTTATGAGGCGCTGGA	-----C	1258
Db	519	AspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGly	-----C	538
Qy	1259	AAAGCCATATGATTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGAGTCTCTGAATCG	-----C	1318
Db	539	LysProTyAspLeu-----GlnGluArgHisSerIleArgValProGluSer	-----C	554
Qy	1319	GGAGAACATTATGAATGCGATCTTTTGGCATCTCTTCAAGAAAAAACCCTTGGATCAGTATT	-----C	1378
Db	555	GlyGluHisTyGluLeuHisLeuLeuHisTyLeuGlnGlnLeuGlySerArgile	-----C	574
Qy	1379	GCTGCTCTAAAGTGATA	1396	
Db	575	AlaAlaLeuLysValIle	580	

## RESULT 12

US-11-140-224-2  
/ Sequence 2: Application US/11140224  
/ Publication No: US20050227280A1  
/ GENERAL INFORMATION:  
/ APPLICANT: INCYTE GENOMICS, INC.  
/ APPLICANT: DELEGANE, Angelo M.; LAL, Preeti G.  
/ APPLICANT: HAPALIA, April J.A.; PATTERSON, Chandra  
/ APPLICANT: WALIA, Narinder K.; KEARNEY, Liam  
/ APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.  
/ APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.  
/ APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.  
/ APPLICANT: NGUYEN, Darniel B.; GANDHI, Ameena R.  
/ APPLICANT: YANG, Junming; HERNANDEZ, Roberto  
/ APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.  
/ APPLICANT: REDDY, Roopa M.; YUE, Henry  
/ APPLICANT: TANG, Y. Tom  
/ TITLE OF INVENTION: PROTEASES  
/ FILE REFERENCE: PI-0085 USN  
/ CURRENT APPLICATION NUMBER: US/11/140,224  
/ CURRENT FILING DATE: 2005-05-31  
/ PRIOR APPLICATION NUMBER: US/10/275,505  
/ PRIOR FILING DATE: 2002-11-04  
/ PRIOR APPLICATION NUMBER: PCT/US01/14651  
/ PRIOR FILING DATE: 2001-05-04  
/ PRIOR APPLICATION NUMBER: 60/209,402  
/ PRIOR FILING DATE: 2000-06-01  
/ PRIOR APPLICATION NUMBER: 60/207,477  
/ PRIOR FILING DATE: 2000-05-25  
/ PRIOR APPLICATION NUMBER: 60/205,803  
/ PRIOR FILING DATE: 2000-05-17  
/ PRIOR APPLICATION NUMBER: 60/203,566  
/ PRIOR FILING DATE: 2000-05-17  
/ PRIOR APPLICATION NUMBER: 60/202,082  
/ PRIOR FILING DATE: 2000-05-04  
/ NUMBER OF SEQ ID NOS: 28  
/ SOFTWARE: PERL Program  
/ SEQ ID NO 2  
/ LENGTH: 580  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: misc feature  
/ OTHER INFORMATION: Incyte ID No: 376067CD1  
US-11-140-224-2

Alignment Scores:  
Pred. No.: 3,02e-221 Length: 580  
Score: 2244.50 Matches: 439  
Percent Similarity: 84.0% Conservative: 3  
Best Local Similarity: 83.5% Mismatches: 5  
Query Match: 75.9% Indels: 79  
DB: 6 Gaps: 6

US-10-825-632-6 (1-1669) x US-11-140-224-2 (1-580)

Qy	2	ACAGGTACAGCAAAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGCTGAA	61
Db	73	ThrGlyThrAlaasnProlysValThrPheLysMetSerGluLeuMetIleaspAlaGlu	92
Qy	62	GCNAGCATCATAGATGTCATAGATAAGCACTATTCACCTTTTCAGATTCTATTGAA	121
Db	93	GlyArgIleIleaspValIleasnPlysGluLeuIleGlnProPheGluLeuPheGlu	112
Qy	122	GGAGTTGAATATATTGCCAGACGCTGGATGGACTCCTCAGGGAAAAATATGCTTGGTCCATC	181
Db	113	GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle	132
Qy	182	CTACTAGATCGCTCCAGACGCTCGGCTACAGATAGTGTGATCTCACTGAATATTATATC	241
Db	133	LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle	152

Qy	242	CCAGTAGAACATGATCTTATGGAAGGACAGACATCATTGAGTCAGTCCCTGATCTCTGTG	301
Db	153	ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal	172
Qy	302	ACGCCACTAATTATCTATGAAGAAACAACAGACATCTCGATAAATATCCATGACATCTTT	361
Db	173	ThrProLeuIleIleTyrGluGluThrThrAspIleIleTrpIleAsnIleHisAspIlePhe	192
Qy	362	CATGTTTTTCCCAAAAGTCACGAAGAGAAATTTGAGTTTATTTTTTGCCTCTCGAATGCAAA	421
Db	193	HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys	212
Qy	422	ACAGGTTTCGGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAACA	481
Db	213	ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg	232
Qy	482	TCCAGTGTGGGCTGCTCTCAAGTGATTTCAAAGTGTCTCTATCAAAGAGGAGATAGCA	541
Db	233	SerSerGlyGlyLeuProAlaProThr	241
Qy	542	ATTACACAGTGGTGAATGGGAAGTT-----CTTGGC-----	571
Db	242	ValThr-----TrpMetIleThrPheMetArgSerLeuGlyThrProSerCysMet	258
Qy	572	-----CGCATGGATCTAATATCCAACTTGATGAAGTCAGAAGCGCTGATATATTTGAA	625
Db	259	CysValThrHisIleValGluIleGlnValAspGluValArgLeuValTyrPheGlu	278
Qy	626	GGCACCAGAAAGTCCCTTTAGAGCATCACCTGACGTACGTACGTAAATCTCTGGA	685
Db	279	GlyThrLysAspSerProLeuGluHisLeuTyrValSerTyrValAsnProGly	298
Qy	686	GAGGTGACAAGGCTGACTACCGTGGCTACTCACATTCTTCTGCTGATCAGTCACACTGT	745
Db	299	GluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGlnHisCys	318
Qy	746	GACTTCTTTATAGTAAGTATAGTAACCCAGAGAAATCCACACTGTGTGTCCTTTACAAG	805
Db	319	AspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyrLys	338
Qy	806	CTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTG	865
Db	339	LeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrIleLeu	358
Qy	866	GATTCTACTGTGCTGTTTATATATGCTCTCCAGAAATTTTCTCTTTTGAAGTACT	925
Db	359	AspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSerThr	378
Qy	926	ACTGGATTATACATTGTTGGGATGCTCTACAAGCTCATGATCTACAGCCTGGAAACAAA	985
Db	379	ThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGlyLysLys	398
Qy	986	TATCCTACTGTGCTGTTTATATATGCTGCTCTCAG-----	1021
Db	399	TyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPhe	418
Qy	1021	-----	1021
Db	419	LysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValValVal	438
Qy	1021	-----	1021
Db	439	IleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLys	458
Qy	1022	---GTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG	1078
Db	459	MetValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThr	478
Qy	1079	GAAAGTTATATGGTCAACCTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCC	1138
Db	479	GluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAla	498
Qy	1139	ATGCAAGACAGAAAAGTTCCCTCTGAACCAAACTGTTTACTGCTCTTACATGGTTCCTG	1198



```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-7

Alignment Scores:
Pred. No.: 1,1e-191 Length: 690
Score: 1957.50 Matches: 372
Percent Similarity: 99.5% Conservative: 1
Best Local Similarity: 99.2% Mismatches: 0
Query Match: 66.2% Indels: 2
DB: 5 Gaps: 1

US-10-825-632-6 (1-1669) x US-10-982-512-7 (1-690)

QY 2 ACAGGTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProlyValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337
QY 62 GGNAGGATCATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAGATTCTATTGAA 121
Db 338 GlyArgIleIleAspValIleAspPheGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GCAGTTGATATATTTGCCAGAGCTGGATGGACTCTCGAGGAAATATCTTTGGTCCATC 181
Db 358 GlyValGluTyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTATAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCACTGAATATTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAGTAGAAGATGATTTATGGAAGGACAGACTCATTTGAGTCAGTGCCTGATTTCTGTG 301
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
QY 302 ACGCCTCAATATCTATGAGAAACAAAGACATCTGGATAAATATCCATGACATCTTT 361
Db 418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437
QY 362 CATGTTTTCCTCCCAAGTACGAGAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAA 421
Db 438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysIys 457
QY 422 ACAGTTTTCGTCATTTATACAAATTTACATCTATTTAAAGGAAGCAATATAACGA 481
Db 458 ThrGlyPheArgHisLeuTyrIleThrSerIleLeuLysGluSerIleTyrLysArg 477
QY 482 TCCAGTGTGGCTGCTGCTCCAGTGTATTTCAAGTGTCTTATCAAGAGGAGATAGCA 541
Db 478 SerSerGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497
QY 542 ATTACCAAGTGTGAATGGAGTTCTTGGCCGGCATGGATCTATATATCCAGTTGATGAA 601
Db 498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517
QY 602 GTCAGAGGCTGTATATTTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTAC 661
Db 518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 537
QY 662 GTAGTCAGTTACTAAATCTCGAGAGGTGCAAGGGCTGACTACCGTGGCTACTACAT 721
Db 538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557
QY 722 TCTTGTGCATGATCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACCAAGAGAAT 781
Db 558 SerCysCysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsn 577
QY 782 CCACACTGTGTGCTCCCTTTACAAAGCTATCAAGTCTCGAAGATACCCAACTTCCAAACA 841
Db 578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597
```

---

```
QY 842 AAGGAATTTGGGCCACCACTTTTGGATTTCAGAGTCTCTCTCTGACTACTCTCCA 901
Db 598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617
QY 902 GAAATTTCTCTTTTCAAGTACTACTGGATTACATTGTATGGGATGCTCTACAAGCCT 961
Db 618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637
QY 962 CATGATCTACAGCTGGAAAGAAATATCTACTGTGTGTTTCATATATGGTGGTCTCAG 1021
Db 638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--Arg 656
QY 1022 GTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAA 1081
Db 657 -LeuLeuLeuLeuGlyProGlnSerLeuCysGlySerSerMetIleGlnAspThrArgAs 676
QY 1082 CGTTATATGGTGCACCTGACCAGAAATGAACAGGGCTATTACT 1124
Db 676 nValIleTrpValThrLeuThrArgMetAsnArgAlaIleThr 690

RESULT 15
US-09-976-674-19
; Sequence 19, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Q4, Steve
; APPLICANT: Akisanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junich, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 0869
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIORITY FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIORITY FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19

Alignment Scores:
Pred. No.: 2,59e-176 Length: 658
Score: 1808.00 Matches: 338
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 61.1% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-6 (1-1669) x US-09-976-674-19 (1-658)

QY 2 ACAGGTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATGCTGAA 61
Db 318 ThrGlyThrAlaAsnProlyValThrPheLysMetSerGluLeuMetIleAspAlaGlu 337
QY 62 GGAAGGATCATAGATGTCATAGATAAGGACTTAATTAACCTTTGAGATTCTATTGAA 121
Db 338 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 357
QY 122 GGAGTTGAATATATTCGAGAGCTGGATGCACTCTCGAGGAAATATGCTTGTGTCATC 181
Db 358 GlyValGluTyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIle 377
QY 182 CTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAATATTATTC 241
Db 378 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 397
QY 242 CCAATGAGATGATGTTATGGAAGGACAGACTCATTTAGTCACTGCTGATTTCTGTG 301
Db 398 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 417
```

Qy 302 AGCCCACTAATTTATCTATGAAGAACACACAGACATCTGGATAAATATCCATGACATCTTT 361  
Db |||||||  
418 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 437  
Qy 362 CATGTTTTCCCAAGTCACGAGAGAAATTCAGTTTATTTTGGCTCTGAATGCAAA 421  
Db |||||||  
438 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 457  
Qy 422 ACAGGTTTCCGTCATTATACAAAATACATCTATTTAAAGGAAGCAAAATATAACGA 481  
Db |||||||  
458 ThrGlyPheArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArg 477  
Qy 482 TCCAGTGTGGCTGCTGCTCCAAAGTCATTTCGAAGTCTCTATCAAAAGAGAGATAGCA 541  
Db |||||||  
478 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 497  
Qy 542 ATTACAGTGGTCAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGA 601  
Db |||||||  
498 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 517  
Qy 602 GTCAGAAAGCTGCTATATTTGAAGGCACCAAGACTCCCTTTAGACATCACCTGTGTAC 661  
Db |||||||  
518 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 537  
Qy 662 GTAGTCAGTTACGTAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACAT 721  
Db |||||||  
538 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 557  
Qy 722 TCTTGCTGCATCAGTCAGCAGTGTGACTTCTTTATAGTAAGTATAGTACCAAGAGAAT 781  
Db |||||||  
558 SerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 577  
Qy 782 CGACACTGTGTGCTCTTTACAAAGCTATCAAGTCTGAGATGACCCAACTTGCAAAACA 841  
Db |||||||  
578 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 597  
Qy 842 AAGGAATTTGGGCCCACTTTTGGATTTCAGCAGGTCTCTTCTGACTATCTCTCTCCA 901  
Db |||||||  
598 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 617  
Qy 902 GAAATTTCTCTTTGAAAGTACTACTGATTTACATTTGATGGGATGCTCTACAGCCT 961  
Db |||||||  
618 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 637  
Qy 962 CATGATCTACAGCTCGAAAGAAATATCTCTACTGTGCTGTTCATATATGTTGGT 1015  
Db |||||||  
638 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655

Search completed: May 2, 2006, 03:08:15  
Job time : 277.002 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 02:50:19 ; Search time 5.57803 Seconds  
(without alignments)  
2721.092 Million cell updates/sec

Title: US-10-825-632-6  
Perfect score: 2957  
Sequence: 1 aacaggtacagcaatccta.....aaaaaaaaaaaaaaaa 1669

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 464238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DRV=xlp  
-Q=/abss/ABSSWEB\_spool/US10825632/runat\_01052006\_110000\_3533/app\_query.fasta.1  
-DB=Published Applications\_AA\_New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abs02p  
-USER=US10825632 @CGN\_1\_1\_59 runat\_01052006\_110000\_3533 -NCPU=6 -ICPU=3  
-NO\_WMAP -NEG\_SCORES=0 -WAIT -DSPELCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6 -FGAEXT=7  
-YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_AA\_New.\*

- 1: /SID55/ptodata/1/pubpaa/US08\_NEW\_PUB.pepl.\*
- 2: /SID55/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /SID55/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /SID55/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /SID55/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 6: /SID55/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 7: /SID55/ptodata/1/pubpaa/US09\_NEW\_PUB.pepl.\*
- 8: /SID55/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 9: /SID55/ptodata/1/pubpaa/US10\_NEW\_PUB.pepl.\*
- 10: /SID55/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 11: /SID55/ptodata/1/pubpaa/US11\_NEW\_PUB.pepl.\*
- 12: /SID55/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2422	81.9	882	11	US-11-151-601-20
2	1298.5	43.8	497	11	US-11-264-096-1594
3	1294.5	43.8	497	11	US-11-264-096-1593
4	588.5	19.9	349	11	US-11-264-096-1591
5	577	19.5	109	11	US-11-176-951-10

6	394	13.3	99	11	US-11-176-951-11	Sequence 11, Appl
7	359.5	12.2	745	11	US-11-079-463-6408	Sequence 6408, Ap
8	269.5	9.1	627	11	US-11-079-463-7758	Sequence 7758, Ap
9	238	8.0	738	11	US-11-208-288-4	Sequence 4, Appli
10	238	8.0	766	9	US-10-501-035-234	Sequence 234, App
11	238	8.0	766	11	US-11-208-288-2	Sequence 2, Appli
12	234	7.9	762	11	US-11-116-939-13	Sequence 13, Appl
13	234	7.9	766	9	US-10-522-789-2	Sequence 2, Appli
14	229	7.7	760	11	US-11-208-288-6	Sequence 6, Appli
15	214	7.2	760	11	US-11-186-284-55	Sequence 55, Appl
16	145	4.9	102	11	US-11-176-951-7	Sequence 7, Appli
17	139	4.7	115	11	US-11-176-951-9	Sequence 9, Appli
18	139	4.7	115	11	US-11-176-951-12	Sequence 12, Appl
19	133.5	4.5	102	11	US-11-176-951-8	Sequence 8, Appli
20	111.5	3.8	1019	9	US-10-995-561-982	Sequence 982, App
21	109.5	3.7	959	9	US-10-467-962B-4	Sequence 4, Appli
22	104.5	3.5	352	11	US-11-098-686-11006	Sequence 11006, A
23	104	3.5	468	11	US-11-156-084-250	Sequence 250, App
24	102	3.4	732	9	US-10-518-599-22	Sequence 22, Appl
25	98.5	3.3	623	11	US-11-188-298-16754	Sequence 16754, A
26	97	3.3	624	11	US-11-079-463-7504	Sequence 7504, Ap
27	95.5	3.2	570	11	US-11-143-984A-10	Sequence 10, Appl
28	95.5	3.2	624	11	US-11-079-463-10187	Sequence 10187, A
29	94.5	3.2	2323	9	US-10-793-626-760	Sequence 760, App
30	93	3.1	737	11	US-11-079-463-9281	Sequence 9281, Ap
31	92.5	3.1	421	11	US-11-045-004-2837	Sequence 2837, Ap
32	92.5	3.1	595	11	US-11-079-463-8616	Sequence 8616, Ap
33	92.5	3.1	2725	11	US-11-113-424-52	Sequence 52, Appl
34	91.5	3.1	1461	11	US-11-183-136-28	Sequence 28, Appl
35	91	3.1	439	11	US-11-075-185-11	Sequence 11, Appl
36	90.5	3.1	328	11	US-11-188-298-20669	Sequence 20669, A
37	90.5	3.1	487	9	US-10-467-657-2272	Sequence 2272, Ap
38	89.5	3.0	486	11	US-11-188-298-21031	Sequence 21031, A
39	89.5	3.0	3580	9	US-10-510-941-14	Sequence 14, Appl
40	89	3.0	563	11	US-11-079-463-9616	Sequence 9616, Ap
41	89	3.0	622	11	US-11-188-298-15979	Sequence 15979, A
42	89	3.0	628	11	US-11-087-099-4659	Sequence 4659, Ap
43	89	3.0	828	9	US-10-995-561-983	Sequence 983, App
44	89	3.0	918	9	US-10-995-561-981	Sequence 981, App
45	88.5	3.0	422	9	US-10-517-939-238	Sequence 238, App

ALIGNMENTS

RESULT 1  
US-11-151-601-20  
; Sequence 20, Application US/11151601  
; Publication No. US20060003413A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Weich, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE, AND PROTEASE FAMILY MEMBERS AND USES THEREOF  
; FILE REFERENCE: MF100-054PFCFLOMIDVIM  
; CURRENT APPLICATION NUMBER: US/11/151,601  
; CURRENT FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: US 10/170,789  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/882,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15



RESULT 2  
US-11-264-096-1594  
; Sequence 1594, Application US/11264096



```
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1593
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (183)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-264-096-1593

Alignment Scores:
Pred. No.: 9,52e-117 Length: 497
Score: 1294.50 Matches: 247
Percent Similarity: 61.2% Conservative: 57
Best Local Similarity: 49.7% Mismatches: 92
Query Match: 43.8% Indels: 101
DB: 11 Gaps: 2

US-10-825-632-6 (1-1669) x US-11-264-096-1593 (1-497)

QY 179 ATCTCTAGATCGCTCCAGACTCGCTCAGACTAGTGTGATCTCACTGATTAATTT 238
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1 MetPheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProAlaLeuPhe 20

QY 239 ATCCAGTAGAAGATGATGTTATGGAAGGAGGAGACTCATTGAGTCAGTCTGATCT 298
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 21 IleProSerThrGluAsnGluGlnArgLeuAlaSerAlaArgAlaValProArgAsn 40

QY 299 GTGAGCCCACTAATATATCTATGAAGAAACACAGACATCTGGATAAATATCATCATC 358
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 41 ValGlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnValHisAspIle 60

QY 359 TTTTCATGTTTTCCTGATTTTATGAGAGAAATTCAGTTTATTTTTCCTCTGAA 415
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 61 PheTyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGlu 80

QY 416 TCAAAACAGGTTTCCCTGATTTATACAAATACATCTATTATTAAGAAAGCAATAT 475
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 81 CysLysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyr 100

QY 476 AAACGATCCAGTGTGGCTGCTGCTCCAGTGATTTCAAGTGCTCTATCAAGAGAG 535
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 101 AspTrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGlu 120

QY 536 ATAGCAATTACAGTGTGAATGGAGTTCTTGGCCGCGCATGATCTAATATCAAGTT 595
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 121 IleAlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpVal 140

QY 596 GATGAAGTCAGAGGCTGTATATTGTAAGGACCAAGACTCCCTTTAGAGCATCAC 655
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 141 AsnGluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHis 160

QY 656 CTGTACTAGTCAATACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTAC 715
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 161 LeuTyrValValSerTyrGluAlaGlyGluIleValArgLeuThrThrProGlyPhe 180

QY 716 TCACATTTGTCATCAGTCACTGCTGACTTCTTTATAAGTAAGTATAGTAACACAG 775
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 181 SerHis***CysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal 200

QY 776 AAGAAATCCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCTCAAGATGACCCCACTGC 835
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 201 SerThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHis 220

QY 836 AAACAAAGGAATTTTGGGCCACCAATTTTGGATTTCAGGAGTCTCTTCTCTGACTACT 895
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```
Db 221 LysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrVal 240
QY 896 CTCACAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATTTATGGGATGCTCTAC 955
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 241 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyr 260
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 956 AAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTGCTGCTTTCATATATGGTGGT 1015
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 261 LysProHisAlaLeuGlnProGlyLysHisPheThrValLeuPheValTyrGlyGly 280
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 1016 CCTCAG----- 1021
   |||||
Db 281 ProGlnValGlnLeuValAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThr 300
   |||||
QY 1021 ----- 1021
   |||||
Db 301 LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly 320
   |||||
QY 1021 ----- 1021
   |||||
Db 321 LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal 340
   |||||
QY 1021 ----- 1021
   |||||
Db 341 GluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle 360
   |||||
QY 1021 ----- 1021
   |||||
Db 361 HisGlyTrpSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal 380
   |||||
QY 1022 ----- 1022
   |||||
Db 381 PheLysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyr 400
   |||||
QY 1076 AGCAAGCGTTATATGGTTCACCTGACCAAGAACACAGGGCTATTACTTAGGATCTGTG 1135
   |||||
Db 401 ThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerVal 420
   |||||
QY 1136 GCCATCAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGTCTTACATGGTTTC 1195
   |||||
Db 421 AlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleLeuHisGlyPhe 440
   |||||
QY 1196 CTGGATGAGAATGTCCTATTTTGACATPACCAGTATATTACTAGTATTTTAGTGGGCT 1255
   |||||
Db 441 LeuAspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAla 460
   |||||
QY 1256 GGAAGCCCATATGATTTTACAGATCTATCCTCAGAGAGACACAGACATAAGATTCCTGAA 1315
   |||||
Db 461 GlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCysProGlu 480
   |||||
QY 1316 TCGGAGAGAACATTATGAACCTGCATCTTTTGCATCTCTTCAAGAAACCTT 1366
   |||||
Db 481 SerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 497
   |||||
```

```
RESULT 4
US-11-264-096-1591
; Sequence 1591, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS46D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; PRIOR FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1591
; LENGTH: 349
;
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (183)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (191)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (192)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (334)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (344)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (345)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (348)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;
; US-11-264-096-1591
;
Alignment Scores:
Pred. No.: 1,5e-48 Length: 349
Score: 588.50 Matches: 111
Percent Similarity: 72.0% Conservative: 33
Best Local Similarity: 55.5% Mismatches: 55
Query Match: 19.9% Indels: 1
DB: 11 Gaps: 1

US-10-825-632-6 (1-1669) x US-11-264-096-1591 (1-349)
Qy 179 ATCTACTAGATCGCTCCAGACTCGCCTCAGACATAGTGTGATCTCACCCTGAATATATT 238
Db 1 MetPheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPhe 20
Qy 239 ATCCAGTAGAAGATGATGTTATGGAAGGAGGAGACTCATTGAGTCAGTCGCCTGATTCT 298
Db 21 IleProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsn 40
Qy 299 GTGACGCCAATAATTATCTATGAAGAACACAGACATCTGGATAAATATCCATGCATC 358
Db 41 ValGlnProTrpValValTrpGluGluValThrAsnValTrpIleAsnValHisAspIle 60
Qy 359 TTTTCATGTTTTCCCAAGTACAC--GAAGAGGAATGAGTTTATTTTGGCTCTGAA 415
Db 61 PheTyzProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArgAlaAsnGlu 80
Qy 416 TDCAAACAGGTTTCCTCGTCATTATATACAAATTTACATCTATTTTAAAGGAAACAAATAT 475
Db 81 CysLysThrGlyPheCysHisLeuTyzLysValThrAlaValLeuLysSerGlnGlyTyz 100
Qy 476 AAACGATCCAGTGGTGGCTGCTGCTCCAAAGTATTTCAAGTATTTCAAGTCTCTATCAAGAGGAG 535
Db 101 AspTrpSerGluProPheSerProGlyGluAspGluPheLysCysProIleLysGluGlu 120
Qy 536 ATAGCAATACCAAGTGGTGAATGGGAGTTTCTTGGCCGGCATGGATCTAATATCAAGTT 595
Db 121 IleAlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpVal 140
```

```

Qy 596 GATGAAGTCAGAAAGCTGCTATATTTTGAAGCAACCAAGACTCCCTTTAGAGCATCAC 655
Db 141 AsnGluGluThrLysLeuValTyzPheGlnGlyThrLysAspThrProLeuGluHisHis 160
Qy 656 CTGTACGTAGTCAGTTAGTAAATCTCAGAGGTGACAAGGCTGACTGACCGTGGCTAC 715
Db 161 LeuTyzValValSerTyzGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPhe 180
Qy 716 TCACATCTCTGTCATCAGTCAGTCACACTGTGACTTCTTTATAAGTAAGTATAGTAAACGAG 775
Db 181 SerHis***CysSerMetSerGlnAsnPhe*****PheValSerHisIleThrAlaGln 200
;
RESULT 5
US-11-176-951-10
; Sequence 10, Application US/11176951
; Publication No. US2006002431A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 10
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-11-176-951-10
;
Alignment Scores:
Pred. No.: 1,39e-47 Length: 109
Score: 577.00 Matches: 109
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 19.5% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-6 (1-1669) x US-11-176-951-10 (1-109)
Qy 1070 GGATACACGAACGTTATATGGTCACCTGACAGAAATGAACAGGGCTATTACTTAGGA 1129
Db 1 GlyTyzThrGluArgTyzMetGlyHisProAspGlnAsnGluGlnGlyTyzTyzLeuGly 20
Qy 1130 TCTGTGCCATGCAACAGCAAAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACAT 1189
Db 21 SerValAlaMetGlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHis 40
Qy 1190 GGTTCCTGTCAGAGATCTCCATTTGACATACACATATATTACTGAGTTTTTTAGTG 1249
Db 41 GlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuVal 60
Qy 1250 AGGCTGGAAGGCATATCATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTT 1309
Db 61 ArgAlaGlyLysProTyzAspLeuGlnIleTyzProGlnGluArgHisSerIleArgVal 80
Qy 1310 CCTGAATCGGGAGAACATTATGAACATGCTCTTTTGCACACTACCTTCAAGAAAACTTGA 1369
Db 81 ProGluSerGlyGluHisTyzGluLeuHisLeuLeuHisLeuGlnGluAsnLeuGly 100
Qy 1370 TCACGATTCTGCTCTTAAAGTGATA 1396
Db 101 SerArgIleAlaAlaLeuLysValIle 109
;
RESULT 6
US-11-176-951-11
; Sequence 11, Application US/11176951
```

```
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: 08842, 0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 3.3
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-11

Alignment Scores:
Pred. No.: 6,62e-30 Length: 99
Score: 394.00 Matches: 72
Percent Similarity: 82.8% Conservative: 10
Best Local Similarity: 72.7% Mismatches: 17
Query Match: 13.3% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-6 (1-1669) x US-11-176-951-11 (1-99)
QY 1070 GGATACACGAGCGTTATATGGTCACCGTCACGACAGATGAACAGCGCTATTACTTAGGA 1129
DB 1 GlyTyrThrGluArgTyrMetAspValProGluAsnGlnHisGlyTyrGluAlaGly 20
QY 1130 TCTGTGCCATCGACGAGAAAGTCCCTCTCGAACCAATCGTTTACTGCTCTTACAT 1189
DB 21 SerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuLeuHis 40
QY 1190 GGTTCCTCGATGAGAAATGCTCCATTTGCAATACCATGATATTTACTGAGTTTGTAGT 1249
DB 41 GlyPheLeuAspGluAsnValHisPheHisThrAsnPheLeuValSerGlnLeuLeu 60
QY 1250 AGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTT 1309
DB 61 A-GAlaGlyLeuProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCys 80
QY 1310 CTGTAATCGGGAGAACATTATGAACTGCATCTTTTGGCACTACCTTCAAGAAACCTT 1366
DB 81 ProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 99

RESULT 7
US-11-079-463-6408
; Sequence 6408, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Bretton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6408
; LENGTH: 745
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-6408
```

```
Alignment Scores:
Pred. No.: 2,54e-26 Length: 745
Score: 359.50 Matches: 125
Percent Similarity: 36.4% Conservative: 74
Best Local Similarity: 22.9% Mismatches: 163
Query Match: 12.2% Indels: 185
DB: 11 Gaps: 17

US-10-825-632-6 (1-1669) x US-11-079-463-6408 (1-745)
QY 20 AAAGTCACCTTTTAAGATGTCAGAAATAATGATGATGCTGAAGGAAGGATCATAGATGTC 79
DB 286 LysValThr---ArgGlnValLysLeuProIleAspAlaAspGly----- 299
QY 80 ATAGATAAGGAACATAATTCAACCTTTTGAGATTCTATTGAAGGAGTTGAATATATTGCC 139
DB 300 -----TyrIlePro 302
QY 140 AGAGCTGGATGGACTCTCGAGGAAAATATGCTTGTCCATCTACTAGATCGTCCAG 199
DB 303 ArgIleArgPheThrGlnAspProAsnLysLeuAlaIleMetThrLeuAsnArgHisGln 322
QY 200 ACTCGCTACAGATAGTGTGATCTCACCCTGAATATTATTATCCAGTAGAGATGATGTT 259
DB 323 AsnArgPheAspMetTyrPheAlaAspPro----- 332
QY 260 ATGGAAGGCGAGAGACTCATTCAGTCAGTCTGATCTCTGTGAGCCACTAATTATCTAT 319
DB 333 -----ArgSerThrValCysLysLeuAlaLeuArgAspGluSerProTyrTyrIleAsn 350
QY 320 GAGAAACACACAGACATCTGGGATAATATCCATGACATCTTTCATGTTTTCCTCCAAAGT 379
DB 351 Glu-----AsnValPheAspAsnIleGlnPheTyrProGlu--- 362
QY 380 CACGAAGAGAAATTGAGTTTATTTTTCCTCTCGAATGCAAAACAGGTTTCCGTCATTTA 439
DB 363 -----TyrPheSerPheValSerAsp---LysSerGlyTyrProHisLeu 376
QY 440 TACAAAATTACATCTATTATTTAAAGGAAAGCAAAATAAATCAACGATCCAGTGTGGGCTGCCT 499
DB 377 Tyr-----TrpTyrSerMetAsnGlyAsnLeu--- 385
QY 500 GCTCCAAGTGATTTCAAGTGCTCTATCAAGAGAGAGATAGCAATTACAGTGGTGAATGG 559
DB 386 -----IleLysGln-----ValThrSerGlyAsnTyr 394
QY 560 GAAGTTCTTGGCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATAT 619
DB 395 GluValLysAsnPheIleGlyTyrAsnProAspThrAsnGlu-----PheTyr 410
QY 620 TTGTAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGTAAAT 679
DB 411 TyrThrSerAsnGluSerProMetArgGlnAlaValTyrLysIleAsp-----Arg 428
QY 680 CTGGAGAGGTGACAGGCTGACCGGTGCTACTCACAATCTTGTCTGCATCAGTCAG 739
DB 429 LysGlyLysLysMetLysLeuSerAsnGlnProGlyThrAsnSerProIlePheSerSer 448
QY 740 CACTGTGACTCTTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 799
DB 449 SerMetLysTyrPheMetAsnLysPheThrSerLeuAspThrProMetLeuIleThrLeu 468
QY 800 TACAAGCTATCAAGTCTCTGAAGAGTACCCCAACTTGGCAAAACAAAGGAATTTTGGGCCACC 859
DB 469 -----AsnAspAsnThrGlyLysValLysValLysThrLeuValThr 481
QY 860 ATTTGGATTACAGCAGGTCCTCTTCTGACTATATCTCTCCA-----GAAATTTCTCT 913
DB 482 AsnAspLysLeuLysGlnLysLeuAlaGluTyrAlaIleProGlnLysGluPhePheThr 501
QY 914 TTTGAAAGTACTACTGAGTTTACATTTGATGGGATGCTCTACAGGCTCTCATGATCTACAG 973
DB 502 PheLysThrThrGluGlyValAspLeuAsnGlyTyrTrpMetMetLysProValAsnPheAsp 521
```



```
Db 435 TyrMetAlaAsnLysGlyTyrIleMetPheThrValAspGlyArgGlySerSerAsnArg 454
Qy 1018 -----
Db 455 GlyLeuAspPheGluAsnValThrPheArgGlnLeuGlyIleGluGluGlyArgAspGln 474
Qy 1018 -----
Db 475 ValLysGlyThrGluPheLeuLysSerLeuProTyrValAspGlyAsnArgIleGlyVal 494
Qy 1018 -----
Db 495 HisGlyTrpSerPheGlyClyHisMetThrThrAlaLeuLeuLeuArgTyrProGluIle 514
Qy 1019 ---CAGTTGCTATTGCTGGGCCAGTCACTCTGTGGATCTTCTAATGATACAGGATAC 1075
Db 515 PheLysValGlyValAlaGlyGlyProValIleAspTrpGlyTyrGlyValMetTyr 534
Qy 1076 ACGAAGCTTATATGGTCACCTGACCAAGATCAACAGGCTATTACTTTAGGATCTGTG 1135
Db 535 GlyGluArgTyrMetAspThrProGlnSerAsnProLysGlyTyrLysGluCysAsnLeu 554
Qy 1136 GCATGCCAAGCAGAAAGTTCCCTCTGAACCAATGTTTACTGCTTCTTACATGGTTTC 1195
Db 555 LysAsnLeuAlaGlyAsnLeu-----LysGlyHisLeuMetIleHisAspAsp 571
Qy 1196 CTGGATGAGATGTCCATTTGCACATACCATATATTACTGAGTTTCTTGTAGGAGGCT 1255
Db 572 HisAspAspThrCysValProGlnHisThrLeuSerPheMetLysAlaCysIleAspAla 591
Qy 1256 GGAAGGCCATATGATTACAGATCTATCTCAGGAGACACAGCATAAGATCTCTCGAA 1315
Db 592 ArgThrTyrProAspLeuPheIleTyrProCysHisIleAsnVal----- 607
Qy 1316 TCGGAGAA-----CATTATGAAGTCACTTTTGGTACCTTTCAGTCTTCAAGAAACCTT 1366
Db 608 SerGlyArgAspArgValHisLeuHisGluLysIleThrArgTyrPheGluAspTyrLeu 627

RESULT 9
US-11-208-288-4
; Sequence 4: Application US/11208288
; Publication No: US20060051366A1
; GENERAL INFORMATION
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; PRIOR FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-4

Alignment Scores:
Pred. No.: 148-14 Length: 738
Score: 238.00 Matches: 120
Percent Similarity: 33.5% Conservative: 71
Best Local Similarity: 21.1% Mismatches: 193
Query Match: 8.0% Indels: 186
DB: 11 Gaps: 24

US-10-825-632-6 (1-1669) x US-11-208-288-4 (1-738)

Qy 5 GGTACACCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATTCGCTGAAGGA 64
Db 232 GlyAlaValAsnProThrValLysPhePheVal-----ValAsnThrAspSerLeuSer 249
```

```
Qy 65 AGGATCATAGATGCTCATAGTAAGGAACCTAACTTCAACCTTTTGAGATTCTATTGAAGGA 124
Db 250 SerValThrAsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIle---Gly 268
Qy 125 GTTGAATATATATCCAGAGCTGGATGGACTCTCTGAGGAAAAATATGCTTGGTCCATCCTA 184
Db 269 AspHisTyrLeuCysAspValThrTrpAla----- 278
Qy 185 CTAGATGCTCTCCAGACTCGCTCATAGATAGTGTGATCTCACCTGAATATTATTATCCCA 244
Db 279 -----ThrGlnGluArgIleSerLeuGlnTrpLeu----- 288
Qy 245 GTAGAAGATGATGTTATGGAAGCGACAGACTCATTTGAGTCAGTGCCTGATTCGTGACG 304
Db 289 -----ArgArgIleGlnAsnTyr-----SerValMet 297
Qy 305 CCACTAATTATCTATGAAGAAACACACAGCATCTGGATAAATATCCATGACATCTTTCAT 364
Db 298 AspIleCysAspTyrAspGluSerSerGlyArgTrp---AsnCys----- 311
Qy 365 GTTTTCCCAAGTCAAGCAGAGAAATGAGTTTATTTTGGCTCTGAATGCAAAACA 424
Db 312 LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrpValGlyArgPheArgPro 331
Qy 425 GGTTCCTCGTCAT-----TTATACAAAATTACATCTATTTTAAAG 463
Db 332 SerGluProHisPheThrLeuAspGlyAsnSerPheTyrLysIleLeuSer-----Asn 349
Qy 464 GAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAAGTGTCT 523
Db 350 GluGluGlyTyrArgHisIle-----CysTyrPheGlnIleAsp 362
Qy 524 ATCAAGAGAGATAGCAATTAACGATGGTGAATGGGAAGTCTTGGCCGCGATGGATCT 593
Db 363 LysLeuAspCysThrPheIleThrLysGlyThrTrpGluValIleGly----- 378
Qy 584 AATATCCAAAGTTCATGAAGCTGATATATTTTGAAGGCACC-----AAAGACTCC 640
Db 379 ---IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLysGlyMet 397
Qy 641 CTTTATAGAGCATCACCTGATGATGATTAACCTGAGAGAGTGCACAAAGCTG 700
Db 398 ProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyrThrLysValThrCysLeu 417
Qy 701 ACTGACCGTGGCTACTCACATTTCTTGTGATCAGT---CAGCACGTGACTTCTTTATA 757
Db 418 -----SerCysGluLeuAsnProGluArgCysGlnTyrTyrSer 430
Qy 758 AGTAAGTATAGTAACCAAGAAATCCACAC-----TGT-----GTGTCC 796
Db 431 ValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuPro 450
Qy 797 CTTTACAAGCTATCAAGTCTGAAGATGACCAACTTGCAAAACAAGGAATTTGGGCC 856
Db 451 LeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArgValLeuGluAsp---Asn 469
Qy 857 ACCATTTTGGATTGACAGAGGCTCTTCTGCTGATCTATCTCTCCAGAA-----ATT 907
Db 470 SerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLysLysLeuAsp 487
Qy 908 TTCTCTTTTGAATACTACTGGAATTTACATGTTGATGGATGCTCTACAAAGCTCATGAT 967
Db 488 PheIleIleLeuAsnGluThrLysPhe---TrpTyrGlnMetIleLeuProHis--- 505
Qy 968 CTACAGCTGGAAAGAAATATCTCTGCTCTTCTATATATGTTGGTCTCT----- 1018
Db 506 PheAspLysSerLysTyrProLeuLeuAspValTyrAlaGlyProCysSerGln 525
Qy 1018 ----- 1018
Db 526 LysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIle 545
```



Qy	1018	---	1018
Db	546	IleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAla	565
Qy	1018	---	1018
Db	566	IleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe	585
Qy	1018	---	1018
Db	586	SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGly	605
Qy	1019	---	1036
Db	606	TyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaVal	625
Qy	1037	GCCCACTCACTCTGTGATCTCTATGATACAGATACAGGACGTTATATGGT---	1093
Db	626	AlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArgTyrMetGlyLeu	645
Qy	1094	---CACCCCTGACCAGATGAACAGGCTATTACTTAGGATCTGGCCATCGAACGACAA	1150
Db	646	ProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGlu	665
Qy	1151	AGTTCCCTCTGACCAATCGTTTACTCTCTTACATGTTCTCTGATGAGAATGTC	1210
Db	666	AsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAsnVal	683
Qy	1211	CATTTCGACATACAGTATATTACTGAGTCTTTTAGTGAGGCTGGAAAGCCATGAT	1270
Db	684	HisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGln	703
Qy	1271	TTACAGATCTATCTCAGGAGACACACAGCATAGATTCCTGAATCGGGAGACATTAT	1330
Db	704	AlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIle	723
Qy	1331	GAAGTCATCTTTGGCACTACCTTCAAGAA	1360
Db	724	TyrThrHisMetSerHisPheIleLysGln	733
RESULT 10			
US-10-501-035-234			
; Sequence 234, Application US/10501035			
; Publication NO. US20060046249A1			
; GENERAL INFORMATION:			
; APPLICANT: Bristol-Myers Squibb Company			
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING			
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE			
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS			
; FILE REFERENCE: D0185 PCT			
; CURRENT APPLICATION NUMBER: US/10/501.035			
; CURRENT FILING DATE: 2004-07-09			
; PRIOR APPLICATION NUMBER: US 60/350,061			
; PRIOR FILING DATE: 2002-01-18			
; NUMBER OF SEQ ID NOS: 795			
; SOFTWARE: Patentin version 3.2			
; SEQ ID NO 234			
; LENGTH: 766			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-501-035-234			
Alignment Scores:			
Pred. No.:	1,42e-14	Length:	766
Score:	238.00	Matches:	120
Percent Similarity:	33.5%	Conservative:	71
Best Local Similarity:	21.1%	Mismatches:	193
Query Match:	8.0%	Indels:	186
DB:	9	Gaps:	24
US-10-825-632-6 (1-1669) x US-10-501-035-234 (1-766)			
Qy	5	GGTACAGCAATCCTAAAGTCATCTTTAAGATGTCAGAAATAATGATGATGTCGAGGA	64

Db	260	GlyAlaValAsnProThrValLysPheVal-----ValAsnThrAspSerLeuSer	277
Qy	65	AGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGGAGATCTCTATTTGAAGGA	124
Db	278	SerValThrAsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIle---Gly	296
Qy	125	GTTCAATATATTTGCCAGAGCTGGATGGACTCTGAGGAGAAATATGCTTGGTCCATCCTA	184
Db	297	AspHisTyrLeuCysAspValThrTrpAla-----	306
Qy	185	CTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCA	244
Db	307	-----ThrGlnGluArgIleSerLeuGlnTrpLeu-----	316
Qy	245	GTAAGAATGATGTTATGGAAAGGACAGACTCATTTGAGTCTGAGTCTGCTGATCTCTGACG	304
Db	317	-----ArgAlaGlnAsnTyr-----SerValMet	325
Qy	305	CCACTAATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCAT	364
Db	326	AspIleCysAspTyrAspGluSerSerGlyArgTrp---AsnCys-----	339
Qy	365	GTTTTCCTCCCAAGTCAAGAGAGAAATTTGAGTCTTTATTTTGGCTCTGAATGCAAAACA	424
Db	340	LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrpValGlyArgPheArgPro	359
Qy	425	GGTTTCCGTCAT-----TTATACAAATTTACATCTATTTTAAAG	463
Db	360	SerGluProHisPheThrLeuAspGlyAsnSerPheTyrLysIleIleSer-----Asn	377
Qy	464	GAAGCAATATATAAAGATCCAGTCCAGTGGGGCTGCTCCCAAGTATTTCAAGTGTCTCT	523
Db	378	GluGluGlyTyrArgHisIle-----CysTyrPheGlnIleAsp	390
Qy	524	ATCAAGAGGAGATAGCAATTTACCAGTGTCAATGGAGTCTTCTGGCGGCATGATCT	583
Db	391	LysLysAspCysThrPheIleThrLysGlyThrTrpGluValIleGly-----	406
Qy	584	AATATCCAAAGTTGATGAAGTCAGAGGCTGTATATTTTGAAGGCACC---AAAGACTCC	640
Db	407	-----IleGluAlaLeuThrSerAspTyrLeuTyrTrpIleSerAsnGluTyrLysGlyMet	425
Qy	641	CTTTAGAGCATCCTGCTAGTACTGATTAATCTCTGAGAGGTGACAGGCTG	700
Db	426	ProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyrThrLysValThrCysLeu	445
Qy	701	ACTGACCGTGGCTACTCACTTTCTGTCATCAGT---CAGCACTGTGATCTTTTATA	757
Db	446	-----SerCysGluLeuAsnProGluArgCysGlnTyrTrpSer	458
Qy	758	AGTAAGTATAGTAACAGAGATCCACAC-----TGT-----GTGTCC	796
Db	459	ValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuPro	478
Qy	797	CTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCC	856
Db	479	LeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArgValLeuGluAsp---Asn	497
Qy	857	ACCATTTGGATTGACAGAGTCTCTCTGATGATATCTCTCCAGAA-----ATT	907
Db	498	SerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLysLysLeuAsp	515
Qy	908	TTCTCTTTTGAAGTACTACTGATTTACATTTGATGGATGCTCTACAGACCTCATGAT	967
Db	516	PheIleIleLeuAsnGluThrLysPhe---TrpTyrGlnMetIleLeuProProHis---	533
Qy	968	CTACAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCT	1018
Db	534	PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyProCysSerGln	553
Qy	1018	-----	1018

```
Db 554 LysAlaAspThrValPheArgLeuAsnTrpAlaThrTyLeuAlaSerThrGluAsnIle 573
Qy 1018 -----
Db 574 IleValAlaSerPheAspGlyArgGlySerGlyTyLeuGlnGlyAspLysIleMetHisAla 593
Qy 1018 -----
Db 594 IleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe 613
Qy 1018 -----
Db 614 SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyGlyGly 633
Qy 1019 -----CAGGTTGCTATTGCTGG 1036
Db 634 TyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaVal 653
Qy 1037 GCCCAGTCACCTCTGTGATCTTCTNTGATACAGGATACACGGAAGCTTATATGGT--- 1093
Db 654 AlaProValSerArgTrpGluTyTrpAspSerValTyThrGluArgTyMetGlyLeu 673
Qy 1094 ---CACCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGCCATGCAACAGAA 1150
Db 674 ProThrProGluAspAsnLeuAspHisTyArgAsnSerThrValMetSerArgAlaGlu 693
Qy 1151 AAGTTCCTCTGAACCAATCGTTTACTCTTACATGGTTTCCTGGATCAGAAATGTC 1210
Db 694 AsnPhe-----LysGlnValGluTyLeuLeuIleHisGlyThrAlaAspAspAsnVal 711
Qy 1211 CATTTGGACATACACAGTATTACTAGTATTTAGTGAGGCTGGAAAGCCATATGAT 1270
Db 712 HisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGln 731
Qy 1271 TTACAGATCTATCTCAGGAGACACAGCATAGAAGTTCCTGAATCGGAGAACATTAT 1330
Db 732 AlaMetTrpTyThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIle 751
Qy 1331 GAACGTGATCTTTGCACTACCTTCAAGAA 1360
Db 752 TyrThrHisMetSerHisPheIleLysGln 761

RESULT 11
US-11-208-288-2
; Sequence 2, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: ANGIOGENESIS AND INFLAMMATION
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-2

Alignment Scores:
Pred. No.: 1,428-14 Length: 766
Score: 238.00 Matches: 120
Percent Similarity: 33.5% Conservative: 71
Best Local Similarity: 21.1% Mismatches: 193
Query Match: 8.0% Indels: 186
DB: 11 Gaps: 24

US-10-825-632-6 (1-1669) x US-11-208-288-2 (1-766)
```

```
Qy 5 GGTACAGCAAACTCTAAAGTCACCTTTTAAGATGTGAGAAATAATGATTGATGCTGAAGGA 64
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 GlyAlaValAsnProThrValLysPheVal-----ValAsnThrAspSerLeuSer 277
Qy 65 AGGATCATAGATGTCTATAGATAAGAACTAAATCAACCTTTTGAGATTCTTATTTGAAGGA 124
Db ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: |||||
278 SerValThrAsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIle---Gly 296
Qy 125 GTTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCCTA 184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
297 AspHisTyLeuCysAspValThrTrpAla----- 306
Qy 185 CTAGATCGCTCCACAGCTCGCTACAGATGTTGATCTCACCTGAATTATTATATCCCA 244
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
307 -----ThrGlnGluArgIleSerLeuGlnTrpLeu----- 316
Qy 245 GTAGAAGATGATGTTATGGAAGGCAGAGCTATTGAGTCAGTCGCTGATTTCTGTGACG 304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
317 -----ArgArgIleGlnAsnTy-----SerValMet 325
Qy 305 CCACATAATTATCTAAGAAACAACAGACATCTGGATAAATATATCCATGATCTTTTCAT 364
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
326 AspIleCysAspTyArgPheLeuAspGlyAsnSerPheTyLeuIleIleSer-----Asn 377
Qy 464 GAAGCAAAATAAACAAGTCCAGTGGTGGCTCCCTCAAGTGATTTCAAGTGTCTCT 523
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
378 GluGluGlyTyArgHisIle-----CysTyPheGlnIleAsp 390
Qy 524 ATCAAGAGGAGATAGCAATTACCAGTGGTGAATGGAAAGTCTTGGCGGCATCGATCT 583
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
391 LysLysAspCysThrPheIleThrLysGlyThrTrpGluValIleGly----- 406
Qy 584 AATATCCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCACC---AAAGACTCC 640
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
407 ---IleGluAlaLeuThrSerAspTyLeuTyTyIleSerAsnGluTyLysGlyMet 425
Qy 641 CCTTTAGAGCATCACCTGTACGTAGTACGTAGTAAATAATCTCGAGAGAGTGCACAGGCTG 700
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
426 ProGlyGlyArgAsnLeuTyLysIleGlnLeuIleAspTyThrLysValThrCysLeu 445
Qy 701 ACTGACCGTGGCTACTCACATTTCTGTCATCAGT---CAGCACTGTGACTTCTTATA 757
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
446 -----SerCysGluLeuAsnProGluArgCysGlnTyTrpSer 458
Qy 758 AGTAAGTATAGTAACCAAGAAATCCACAC-----TGT-----GTGTCC 796
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
459 ValSerPheSerLysGluAlaLysTyTyTrpGlnLeuArgCysSerGlyProGlyLeuPro 478
Qy 797 CTTTACAAGCTATCAGTCTCCTGAAGATGCCAATCTGCAAAACAAGAAATTTGGGCC 856
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
479 LeuTyThrLeuHisSerSerValAsnAspLysGlyLeuArgValLeuGluAsp---Asn 497
Qy 857 ACCATTTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAA-----ATT 907
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
498 SerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSerLysLysLeuAsp 515
Qy 908 TTCTCTTTTGAAGTACTACGTGATTTACATTGTATGGATGTCTCTACAAGCCTCATGAT 967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
516 PheIleIleLeuAsnGluThrLysPhe---TrpTyTrpGlnMetIleLeuProHis--- 533
Qy 968 CTACAGCCTGGAAGAAATATCTCTGTGTCTATATATATATATATATATATATATATAT 1018
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
534 PheAspLysSerLysLysTyProLeuLeuLeuAspValTyArgAlaGlyProCysSerGln 553
Qy 1018 ----- 1018
```

Db 554 lysAlaAspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIle 573  
 QY 1018 ----- 1018  
 Db 574 IleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspIleMetHisAla 593  
 QY 1018 ----- 1018  
 Db 594 IleAsnArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe 613  
 QY 1018 ----- 1018  
 Db 614 SerLysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGly 633  
 QY 1019 -----CAGTTTCTATTCTGGG 1036  
 Db 634 TyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaVal 653  
 QY 1037 GCGCAGTCACCTGTGGATCTTCTATGATACAGGATACACGGAAGTTTATATGGGT--- 1093  
 Db 654 AlaProValSerArgTrpGluTyrTyrAspSerValTyrThrGluArgTyrMetGlyLeu 673  
 QY 1094 ---CACCTCGACCAAGTAACAGGCTATTCTTAGGATCTGTGGCCATGCAAGCAGAA 1150  
 Db 674 ProThrProGluAspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGlu 693  
 QY 1151 AAGTTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCCTGATGAGATGTC 1210  
 Db 694 AsnPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAsnVal 711  
 QY 1211 CATTTTGCATACACCATGATATATTAGTATTTTGTAGTGGCTGGAAGCCATATGAT 1270  
 Db 712 HisPheGlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGln 731  
 QY 1271 TTACAGATCTATCTTCAGACAGACACAGCATAGAGTTCTCTGAATCGGAGACCAATTAT 1330  
 Db 732 AlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIle 751  
 QY 1331 GAACGTGATCTTTTGCATCTACCTTCAAGAA 1360  
 Db 752 TyrThrHisMetSerHisPheIleLysGln 761

RESULT 12  
 US-11-116-939-13  
 ; Sequence 13, Application US/11116939  
 ; Publication No. US2005026595A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stephen Tomlinson  
 ; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS  
 ; FILE REFERENCE: 19113.011502  
 ; CURRENT APPLICATION NUMBER: US/11/116,939  
 ; CURRENT FILING DATE: 2005-04-28  
 ; PRIOR APPLICATION NUMBER: 60/565,907  
 ; PRIOR FILING DATE: 2004-04-28  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13  
 ; LENGTH: 762  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic  
 ; OTHER INFORMATION: construct  
 US-11-116-939-13

Alignment Scores:  
 Pred. No.: 3,45e-14 Length: 762  
 Score: 234.00 Matches: 120  
 Percent Similarity: 33.1% Conservative: 70  
 Best Local Similarity: 20.9% Mismatches: 190  
 Query Match: 7.9% Indels: 194

DB: 11 Gaps: 25  
 US-10-825-632-6 (1-1669) x US-11-116-939-13 (1-762)  
 QY 5 GGTACAGCAATCTTAAGATCCTTTTAAAGTATGATGTCAGAAATATGATGATGCTGAAGGA 64  
 Db 256 GlyAlaValAsnProThrValLysPhePheVal-----ValAsnThrAspSerLeuSer 273  
 QY 65 AGGATCATAGATGTCATAGATAAGGAACATAATTCAACCTTTTTCAGATTTCTATTGGAAGGA 124  
 Db 274 SerValThrAsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIle---Gly 292  
 QY 125 GTTGAATATATTGCCAGAGCTGGATGAGTCTCCTGAGGAAAAATATGCTGGTCCATCCTA 184  
 Db 293 AspHisTyrLeuCysAspValThrTrpAla----- 302  
 QY 185 CTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTATTATTATCCCA 244  
 Db 303 -----ThrGlnGluArgIleSerLeuGlnTrpLeu----- 312  
 QY 245 GTAAGAATGATGTTTATGAAAGGCAGAGACTCAATTGAGTCAGTGCTGATTCCTGACG 304  
 Db 313 -----ArgArgIleGlnAsnTyr-----SerValMet 321  
 QY 305 CCACATAATTATCTATGAAGAAACAACAGACATCTGGATATAAATATCCATGACATCTTTCAT 364  
 Db 322 AspIleCysAspTyrAspGluSerSerGlyArgTrp---AsnCys----- 335  
 QY 365 GTTTTTCCTCCCAAGTCACGAAGAGAAATTCAGTTTATTTTGCCTCTGAATCAAAACA 424  
 Db 336 LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrpValGlyArgPheArgPro 355  
 QY 425 GGTTCCTCGTCAT-----TTATACAAAATTATCATCTATTTTAAAG 463  
 Db 356 SerGluProHisPheThrLeuAspGlyAsnSerPheTyrLysIleIleSer-----Asn 373  
 QY 464 GAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCCCAAGTATTTCAAGTGTCT 523  
 Db 374 GluGluGlyTyrArgHisIle-----CystTyrPheGlnIleAsp 386  
 QY 524 ATCAAGAGAGATAGACATACAGTGTGTAAGTGGAGTTCTTGGCCGCGCATGGATCT 583  
 Db 387 LysLysAspCysThrPheIleThrLysGlyThrTrpGluValIleGly----- 402  
 QY 584 AATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATATTTGAAGGCACC-----AAAGACTCC 640  
 Db 403 ---IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLysGlyMet 421  
 QY 641 CCTTTAGAGCATCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 700  
 Db 422 ProGlyGlyArgAsnLeuTyrLysIle-----GlnLeu 432  
 QY 701 ACTGACCGTGGCTACTCACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 745  
 Db 433 SerAsp-----TyrThrLysValThrCysLeuSerCysGluLeuAsnProGluArgCys 450  
 QY 746 GACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACAC-----TGT----- 790  
 Db 451 GlnTyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCysSerGly 470  
 QY 791 -----GTGTCTCTTTTACAAGCTATCAAGTCTCTGAAAGTGAAGTGAAGTGAAGTGAAG 844  
 Db 471 ProGlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuValLeu 490  
 QY 845 GAATTTTGGCCACCATTTTGGATTCAGAGCTCTCTCTCTGATATATCTCTCCAGAA 904  
 Db 491 GluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMetProSer 507  
 QY 905 -----ATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTTAC 955  
 Db 508 LysLysLeuAspPheIleIleLeuAsnGluThrLysPhe---TrpTyrGlnMetIleLeu 526  
 QY 956 AAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTGCTGCTTCTCATATATGGTGT 1015







Mon May 8 13:41:56 2006

```
Db      728 TApTySerAspGlnAsnHisGlyLeu-----SerGlyLeuSerThrAsnHis 743
Qy      1328 TATGAACCTGCATCTTTTGGCACTACCTTCAAGAA 1360
Db      744 LeuTyThrHisMetThrHisPheLeuLysGln 754
```

Search completed: May 2, 2006, 03:10:21  
Job time : 54.8941 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:53:29 ; Search time 5398 Seconds  
(without alignments)  
11404.501 Million cell updates/sec

Title: us-10-825-632-8  
Perfect score: 1083  
Sequence: 1 ggaagaagatgccagatcag.....tatgatacaggatacacgga 1083

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 2842172563 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_to.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1083	100.0	1083	6 AR651455	AR651455 Sequence
2	1083	100.0	1083	8 AF221637	AF221637 Homo sapi
3	1055.8	97.5	4309	6 AR631284	AR631284 Sequence
4	1055.8	97.5	4309	6 AX608737	AX608737 Sequence
5	911	84.1	2098	6 CQ721272	CQ721272 Sequence
6	791.2	73.1	2649	6 AX354795	AX354795 Sequence
7	791.2	73.1	2797	6 AR448400	AR448400 Sequence
8	791.2	73.1	3030	8 AY354202	AY354202 Homo sapi
9	791.2	73.1	3106	6 AX342633	AX342633 Sequence
10	791.2	73.1	3120	6 AR651452	AR651452 Sequence
11	791.2	73.1	3127	8 AF221634	AF221634 Homo sapi
12	791.2	73.1	3143	6 AX354793	AX354793 Sequence
13	791.2	73.1	4535	8 BC040203	BC040203 Homo sapi
14	789.6	72.9	2656	8 AY172659	AY172659 Homo sapi
15	789.6	72.9	2671	6 AR631278	AR631278 Sequence
16	789.6	72.9	2671	6 AX608725	AX608725 Sequence
17	779.2	71.9	4829	6 AR631283	AR631283 Sequence
18	779.2	71.9	4829	6 AX608735	AX608735 Sequence

19	669.6	61.8	4799	9 BC043124	BC043124 Mus muscu
20	669.6	61.8	4799	9 BC059222	BC059222 Mus muscu
21	606.6	56.0	753	6 AX524935	AX524935 Sequence
22	581.8	53.7	2510	6 AX338497	AX338497 Sequence
23	581.2	53.7	3125	8 BC030688	BC030688 Homo sapi
24	579.6	53.5	2161	6 BD157001	BD157001 Primer fo
25	579.6	53.5	2161	6 AX878058	AX878058 Sequence
26	579.6	53.5	2161	8 AK027826	AK027826 Homo sapi
27	549.2	50.7	1449	5 CR353491	CR353491 Gallus ga
28	497.2	45.9	4685	6 AR631288	AR631288 Sequence
29	497.2	45.9	4685	6 AX608745	AX608745 Sequence
30	479.2	44.2	4676	6 AR631287	AR631287 Sequence
31	479.2	44.2	4676	6 AX608743	AX608743 Sequence
32	459	42.4	1197	6 AR651453	AR651453 Sequence
33	459	42.4	1197	8 AF221635	AF221635 Homo sapi
34	434.2	40.1	1669	6 AR651454	AR651454 Sequence
35	434.2	40.1	1669	8 AF221636	AF221636 Homo sapi
36	434.2	40.1	2688	6 AX405771	AX405771 Sequence
37	434.2	40.1	2778	8 AK000290	AK000290 Homo sapi
38	434.2	40.1	2842	6 AX405770	AX405770 Sequence
39	428.2	39.5	4523	6 AR631281	AR631281 Sequence
40	428.2	39.5	4523	6 AX608731	AX608731 Sequence
41	385	35.5	183341	14 AC161443	AC161443 Mus muscu
42	347.4	32.1	2676	6 AX524928	AX524928 Sequence
43	344.2	31.8	2546	8 BC000970	BC000970 Homo sapi
44	344.2	31.8	2592	8 AY374518	AY374518 Homo sapi
45	344.2	31.8	2602	8 AY172660	AY172660 Homo sapi

ALIGNMENTS

RESULT 1:  
AR651455 AR651455 1083 bp DNA linear PAT 20-APR-2005  
LOCUS Sequence 8 from patent US 6881564.  
DEFINITION AR651455  
ACCESSION AR651455  
VERSION AR651455.1 GI:62795941  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1083)  
AUTHORS Abbott, C.A. and Gorrell, M.D.  
TITLE Dipeptidyl peptidases  
JOURNAL Patent: US 6881564-A 8 19-APR-2005;  
The University of Sydney; Sydney;  
AUX.  
FEATURES  
source Location/Qualifiers  
1..1083 /organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match	100.0%;	Score 1083;	DB 6;	Length 1083;
Best Local Similarity	100.0%;	Pred. No. 2.4e-256;		
Matches 1083;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAATTTGATAG	60	
Db	1	CGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAATTTGATAG	60	
QY	61	ATATTTCTGGCTATTTGGTGGTGTCCAAAGCTGAAACACTCCAGTGGTGGTAAATTTCT	120	
Db	61	ATATTTCTGGCTATTTGGTGGTGTCCAAAGCTGAAACACTCCAGTGGTGGTAAATTTCT	120	
QY	121	TAGAATTTCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCC	180	
Db	121	TAGAATTTCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCC	180	
QY	181	TATGTTGGAAACAAGGAGGGCAGATTCCGTTATCTCTTAAACAGGTACAGCAATCC	240	
Db	181	TATGTTGGAAACAAGGAGGGCAGATTCCGTTATCTCTTAAACAGGTACAGCAATCC	240	

QY	241	TAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGAAGGAAGGATCATAGTTGA	300
Db	241	TAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGAAGGAAGGATCATAGTTGA	300
QY	301	TGAAGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACT	360
Db	301	TGAAGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACT	360
QY	361	GTAGCTAGTCAGTTACCTTAATCTCGAGAGGTGACAGGCTGACTGACCGTGCGTACTC	420
Db	361	GTAGCTAGTCAGTTACCTTAATCTCGAGAGGTGACAGGCTGACTGACCGTGCGTACTC	420
QY	421	ACATTCCTGCTGATCAGTCAGTCAGTGTGACTCTCTTTTAAGTAAGTATAGTAACCAAGAA	480
Db	421	ACATTCCTGCTGATCAGTCAGTCAGTGTGACTCTCTTTTAAGTAAGTATAGTAACCAAGAA	480
QY	481	GAATCCACA CTGTGTGTCCTTTTCAAGCTATCAAGTCTGTAAGATGACCCAACTTGCAA	540
Db	481	GAATCCACA CTGTGTGTCCTTTTCAAGCTATCAAGTCTGTAAGATGACCCAACTTGCAA	540
QY	541	AACAAAGAAATTTTGGGCCCAACATTTTGGATTGACAGGTCCTCTTCTGACTATATCTCC	600
Db	541	AACAAAGAAATTTTGGGCCCAACATTTTGGATTGACAGGTCCTCTTCTGACTATATCTCC	600
QY	601	TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTCAATGTATGGATGCTCTACAA	660
Db	601	TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTCAATGTATGGATGCTCTACAA	660
QY	661	GCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCC	720
Db	661	GCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTCC	720
QY	721	TCAGGTGCACTGTGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGTTTGAATACCT	780
Db	721	TCAGGTGCACTGTGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGTTTGAATACCT	780
QY	781	AGCCTCTCTAGTTATGTTGTTGCTAGTATGACACAGGGGATCCTGTACCGAGGCT	840
Db	781	AGCCTCTCTAGTTATGTTGTTGCTAGTATGACACAGGGGATCCTGTACCGAGGCT	840
QY	841	TAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATAGATAAATGACGATCAGGTGA	900
Db	841	TAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATAGATAAATGACGATCAGGTGA	900
QY	901	AGGACTCCATATCTAGCTTCTCGATATGATTTCAATGACTAGATCGTGTGGGCATCCA	960
Db	901	AGGACTCCATATCTAGCTTCTCGATATGATTTCAATGACTAGATCGTGTGGGCATCCA	960
QY	961	CGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATCTT	1020
Db	961	CGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATCTT	1020
QY	1021	CAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGPGAACTCTTATGATACAGGATACAC	1080
Db	1021	CAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGPGAACTCTTATGATACAGGATACAC	1080
QY	1081	GGA 1083	
Db	1081	GGA 1083	

RESULT 2  
AF221637  
LOCUS  
DEFINITION Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, partial cds,  
alternatively spliced.  
ACCESSION AF221637  
VERSION AF221637.1 GI:11095193  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 1083)  
Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W.  
and Gorrell,W.D.  
Cloning, expression and chromosomal localization of a novel human  
dipeptidyl peptidase (DPP) IV homolog, DPP8  
Eur. J. Biochem. 267 (20), 6140-6150 (2000)  
11012666  
2 (bases 1 to 1083)  
Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D.  
Direct Submission  
Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver  
Centre, Centenary Institute of Cell Biology and Cancer Medicine,  
Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia  
Location/Qualifiers  
1..1083  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q22"  
/clone="stPBMD3-3-1"  
/cell\_type="PHA stimulated PBMC"  
1..>1083  
/genes="DPP8"  
1..>1083  
/genes="DPP8"  
/notes="serine protease, member of prolyl oligopeptidase S9b  
family belonging to enzyme clan SC; additional splice  
variants of this mRNA are found in GenBank Accession  
Numbers AF221634, AF221635 and AF221636"  
/codon\_start=2  
/product="dipeptidyl peptidase 8"  
/protein\_id="AA029769.1"  
/db\_xref="GI:11095194"  
/translation="BEDARSAGVATFVLQEPDRYGVWCPKAPETTPSGKILRILY  
ENDESEIIHVTSPLETRRADSFRTKGTANPKVTFKMSIMIDAEGRIIVDEV  
KRIVYEGNDKSPLEHLHVSYPNPGVETRLTDGYSHSCCIQHCDFIFSKYSNQ  
NPKVSLYKLSFDEDDPTCKTFEATILDSAGPLDTPPEIFESFSTFTLYLGM  
YKPHDLOPKKYPTVLLFIYGGQVQLVNNRFGVKYFLRLNLASLGVVVVVDNRGSC  
HRLKEPGAPKYMGGQIEIDDQVGLQSLASRYDFIDLRVGIHGWSYGYSLSLMLM  
ORSIDI PRVAIAGAPVTLWIFDYDTGYT"

Query Match	100.0%	Score 1083	DB 8	Length 1083
Best Local Similarity	100.0%	Pred. No. 2.4e-256		
Matches 1083	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	GGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGAAATTTGATAG	60	
Db	1	GGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGAAATTTGATAG	60	
QY	61	ATATTTCTGGCTATTGTTGGTGTCCAAAGAGCTGAAACAACTCCAGTGGTGGTAAAAATTC	120	
Db	61	ATATTTCTGGCTATTGTTGGTGTCCAAAGAGCTGAAACAACTCCAGTGGTGGTAAAAATTC	120	
QY	121	TAGAAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCC	180	
Db	121	TAGAAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCC	180	
QY	181	TATGTTGGAAACAAGGAGGCGCAGATTTCATTCGGTTATCTTAAACAGGTACAGCAATCC	240	
Db	181	TATGTTGGAAACAAGGAGGCGCAGATTTCATTCGGTTATCTTAAACAGGTACAGCAATCC	240	
QY	241	TAAAGTCACCTTTTAAAGTCTCAGAAATATGATGATGCTGCAAGGAGGATCATAGTTGA	300	
Db	241	TAAAGTCACCTTTTAAAGTCTCAGAAATATGATGATGCTGCAAGGAGGATCATAGTTGA	300	
QY	301	TGAAGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACT	360	
Db	301	TGAAGTCAGAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACT	360	

QY	361	GTACGTAGTCAGTTACGTAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTC	420
Db	361	GTACGTAGTCAGTTACGTAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTC	420
QY	421	ACATTTCTGCTGCATCAGTCAGCAGCTGTGACCTTTTAATAGTAAGTATAGTAACAGAA	480
Db	421	ACATTTCTGCTGCATCAGTCAGCAGCTGTGACCTTTTAATAGTAAGTATAGTAACAGAA	480
QY	481	GAATCCACACGTGTGTCCTTTTACAGCTATCAAGTCTGAGATGACCCAACTTGCAA	540
Db	481	GAATCCACACGTGTGTCCTTTTACAGCTATCAAGTCTGAGATGACCCAACTTGCAA	540
QY	541	AACAAAGGAATTTTGGGCGCCATTTTGGATTGAGAGGTGCTCTTCTGACTATACCTCC	600
Db	541	AACAAAGGAATTTTGGGCGCCATTTTGGATTGAGAGGTGCTCTTCTGACTATACCTCC	600
QY	601	TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAA	660
Db	601	TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAA	660
QY	661	GCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTCAATATATGGTGTCC	720
Db	661	GCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTCAATATATGGTGTCC	720
QY	721	TCAGGTGAGTTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCCGTTGAATACCT	780
Db	721	TCAGGTGAGTTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCCGTTGAATACCT	780
QY	781	AGCCTCTAGGTTATGTGTTGATGATAGACACAGGGGATCTGTCAACGAGGCT	840
Db	781	AGCCTCTAGGTTATGTGTTGATGATAGACACAGGGGATCTGTCAACGAGGCT	840
QY	841	TAAATTTGAAGCGCCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGTGGA	900
Db	841	TAAATTTGAAGCGCCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGTGGA	900
QY	901	AGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGGGCAATCCA	960
Db	901	AGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTGGGCAATCCA	960
QY	961	CGGCTGGTCTTATGGAGATACCTCTCCCTGATGGCAATTAATGACAGGTCAGATATCTT	1020
Db	961	CGGCTGGTCTTATGGAGATACCTCTCCCTGATGGCAATTAATGACAGGTCAGATATCTT	1020
QY	1021	CAGGGTTCATATTCCTGGGCGCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAC	1080
Db	1021	CAGGGTTCATATTCCTGGGCGCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAC	1080
QY	1081	GGA 1083	
Db	1081	GGA 1083	
RESULT 3			
AR631284			
DEFINITION			
Sequence 14 from patent US 6844180.			
ACCESSION			
AR631284			
VERSION			
AR631284.1 GI:59770929			
KEYWORDS			
Unknown.			
ORGANISM			
Unclassified.			
REFERENCE			
1 (bases 1 to 4309)			
AUTHORS			
O.I.S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.			
TITLE			
Serine protease genes related to DPP1V			
JOURNAL			
Patent: US 6844180-A 14 18-JAN-2005;			
Ferring BV;;			
NLX;			
FEATURES			
source			
1..4309			
/organism="unknown"			
/mol_type="genomic DNA"			
Location/Qualifiers			
1..4309			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Query Match 97.5%; Score 1055.8; DB 6; Length 4309;			
Best Local Similarity 99.5%; Pred. No. 1.le-249;			
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;			
QY	1	GGAGAAGATGCCAGATCAGCTGGAGTCTGCTACCTTTGTTCTCTCAAGAAGAAATTTGATAG	60
Db	942	GGAGAAGATGCCAGATCAGCTGGAGTCTGCTACCTTTGTTCTCTCAAGAAGAAATTTGATAG	1001
QY	61	ATATTTCTGGCTATTTGGTGGTCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCT	120
Db	1002	ATATTTCTGGCTATTTGGTGGTCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCT	1061
QY	121	TAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCC	180
Db	1062	TAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCC	1121
QY	181	TATGTTGAAAACAAAGGAGGCGAGATTTCATTCGGTTATCTCTAAACAGGTACAGCAATCC	240
Db	1122	TATGTTGAAAACAAAGGAGGCGAGATTTCATTCGGTTATCTCTAAACAGGTACAGCAATCC	1181
QY	241	TAAAGTCACTTTTAAAGATGTGAGAAATAATGATTTGATGCTGAAGAGAGGATCATAGTTCA	300
Db	1182	TAAAGTCACTTTTAAAGATGTGAGAAATAATGATTTGATGCTGAAGAGGAA-GATCCAAGTTGA	1240
QY	301	TGAAGTCAAGAGGCTGGTATATTTTGAAGGCGACCAAGACTCCCTTTTAGAGCATCACCT	360
Db	1241	TGAAGTCAAGAGGCTGGTATATTTTGAAGGCGACCAAGACTCCCTTTTAGAGCATCACCT	1300
QY	361	GTACGTAGTCAGTTACGTAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGTACTC	420
Db	1301	GTACGTAGTCAGTTACGTAATCCTGGAGAGGTGACAGGCTGACTGACCGTGGTACTC	1360
QY	421	ACATTTCTGCTGCATCAGTCAGCAGCTGTGACTTCTTTTAAAGTAAAGTATAGTAACAGAA	480
Db	1361	ACATTTCTGCTGCATCAGTCAGCAGCTGTGACTTCTTTTAAAGTAAAGTATAGTAACAGAA	1420
QY	481	GAATCCACACGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA	540
Db	1421	GAATCCACACGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAA	1480
QY	541	AACAAAGGAATTTTGGGCGCCATTTTGGATTTGACAGAGTCTCTCTTCTGACTATACTCC	600
Db	1481	AACAAAGGAATTTTGGGCGCCATTTTGGATTTGACAGAGTCTCTCTTCTGACTATACTCC	1540
QY	601	TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAA	660
Db	1541	TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAA	1600
QY	661	GCCTCATGATCTACAGCTGGAAGAAATATCTCTAGTGTCTTCTATATATGTTGGTGG--T	718
Db	1601	GCCTCATGATCTACAGCTGGAAGAAATATCTCTAGTGTCTTCTATATATGTTGGTGGTCT	1660
QY	719	CCTCAGGTGCAAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATACC	778
Db	1661	CCTCAGGTGCAAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATACC	1720
QY	779	CTAGCCTCTCTAGGTTATGTTGGTGTAGTATGATAGACACAGGGGATCTCTGACCCGAGG	838
Db	1721	CTAGCCTCTCTAGGTTATGTTGGTGTAGTATGATAGACACAGGGGATCTCTGACCCGAGG	1780
QY	839	CTTAAATTTGAAGCGCCCTTTAAATATAAATGGGTCAAAATAGAAATTCACGATCAGGTG	898
Db	1781	CTTAAATTTGAAGCGCCCTTTAAATATAAATGGGTCAAAATAGAAATTCACGATCAGGTG	1840
QY	899	GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGGGCATC	958
Db	1841	GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGGGCATC	1900
QY	959	CACGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGAGAGGTGAGATATC	1018
Db	1901	CACGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGAGAGGTGAGATATC	1960

```
QY 1019 TTCAGGTTGCTATTGCTGGGGCCCAAGTCACCTCTGTGGATTTCTATGATACAGGATAC 1078
Db 1961 TTCAGGTTGCTATTGCTGGGGCCCAAGTCACCTCTGTGGATTTCTATGATACAGGATAC 2020

QY 1079 ACGGA 1083
Db 2021 ACGGA 2025

RESULT 4
AX608737
LOCUS AX608737
DEFINITION Sequence 14 from Patent WO0231134.
ACCESSION AX608737
VERSION AX608737.1 GI:28404304
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
TITLE Novel serine protease genes related to dppiv
JOURNAL Patent: WO 0231134-A 14 18-APR-2002;
Ferring BV (NL)
FEATURES
source
1 4309
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 97.5%; Score 1055.8; DB 6; Length 4309;
Best Local Similarity 99.5%; Pred. No. 1.1e-249;
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTGTATG 60
Db 942 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTGTATG 1001

QY 61 ATATTCCTGCTATTGGTGGTGTCCAAAAGCTGAAACAACCTCCAGTGGTGGTAAATTC 120
Db 1002 ATATTCCTGCTATTGGTGGTGTCCAAAAGCTGAAACAACCTCCAGTGGTGGTAAATTC 1061

QY 121 TAGAATCTTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCC 180
Db 1062 TAGAATCTTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCC 1121

QY 181 TATGTTGAAACAGAGGAGGCAGATTCATTCGTTATCTTAAACAGGTACAGCAAAATCC 240
Db 1122 TATGTTGAAACAGAGGAGGCAGATTCATTCGTTATCTTAAACAGGTACAGCAAAATCC 1181

QY 241 TAAAGTCACCTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAAGATCATAGTTGA 300
Db 1182 TAAAGTCACCTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAAGATCATAGTTGA 1240

QY 301 TGAAGTCAGAGGCTGTATATTTGAAGGCCCAAGACTCCCTTTAGAGCATCACCT 360
Db 1241 TGAAGTCAGAGGCTGTATATTTGAAGGCCCAAGACTCCCTTTAGAGCATCACCT 1300

QY 361 GTAGTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACTGACCTGGCTACTC 420
Db 1301 GTAGTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACTGACCTGGCTACTC 1360

QY 421 ACATCTTGCTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTACCCAGAA 480
Db 1361 ACATCTTGCTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTACCCAGAA 1420

QY 481 GAATCCCACTGTGTGCTCTTTTACAGCTATCAAGCTCTGAAGATGACCCAACTTGCAA 540
Db 1421 GAATCCCACTGTGTGCTCTTTTACAGCTATCAAGCTCTGAAGATGACCCAACTTGCAA 1480
```

```
QY 541 AACAAAGGAATTTTGGGCCCAACATTTTGGATTCCAGCAGGTCTCTTCTGACTATACTCC 600
Db 1481 AACAAAGGAATTTTGGGCCCAACATTTTGGATTCCAGCAGGTCTCTTCTGACTATACTCC 1540

QY 601 TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGGATGCTCTACAA 660
Db 1541 TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGGATGCTCTACAA 1600

QY 661 GCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGCTGTTTCATATATGGTGG--T 718
Db 1601 GCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGCTGTTTCATATATGGTGGTCT 1660

QY 719 COTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGAGGAGTCAAGTATTTCCGCTTGAATACC 778
Db 1661 COTCAGGTGCAAGTTGGTGAATATCGGTTTAAAGAGGAGTCAAGTATTTCCGCTTGAATACC 1720

QY 779 CTAGCCTCTCTAGGTTTATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTCCCGAGGG 838
Db 1721 CTAGCCTCTCTAGGTTTATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTCCCGAGGG 1780

QY 839 CTTAAATTTGAAGGCGCTTTTAAATATATAATGGGTCAAATAGAAATTCACGATCAGGTG 898
Db 1781 CTTAAATTTGAAGGCGCTTTTAAATATATAATGGGTCAAATAGAAATTCACGATCAGGTG 1840

QY 899 GAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATGACTTAGATCGTGGGCATC 958
Db 1841 GAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATGACTTAGATCGTGGGCATC 1900

QY 959 CACGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTAATGCGAGGTCAGATATC 1018
Db 1901 CACGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTAATGCGAGGTCAGATATC 1960

QY 1019 TTCAGGTTGCTATTGCTGGGGCCCAAGTCACCTCTGTGGATCTTCTATGATACAGGATAC 1078
Db 1961 TTCAGGTTGCTATTGCTGGGGCCCAAGTCACCTCTGTGGATCTTCTATGATACAGGATAC 2020

QY 1079 ACGGA 1083
Db 2021 ACGGA 2025

RESULT 5
CO721272
LOCUS CO721272
DEFINITION Sequence 7206 from Patent WO02068579.
ACCESSION CO721272
VERSION CO721272.1 GI:42282129
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 7206 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1..2098
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 84.1%; Score 911; DB 6; Length 2098;
Best Local Similarity 88.9%; Pred. No. 6.2e-214;
Matches 1048; Conservative 0; Mismatches 35; Indels 96; Gaps 2;

QY 1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTTGATAG 60
```

Db	835	GGAGAGATGCCAGATCAGCTGGAGTGGCTACTCTTTGTTCTCCAGAGAAATTTGATAG	894
Qy	61	ATAATTCGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTC	120
Db	895	ATAATTCGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTC	954
Qy	121	TAGAAATTCATATGAAGAATGATGAATCTGAGGTGGAAATTAATTCATGTTACATCCC	180
Db	955	TAGAAATTCATATGAAGAATGATGAATCTGAGGTGGAAATTAATTCATGTTACATCCC	1014
Qy	181	TATGTTGGAACAAGGAGGAGATTCATTCCTGTTATCTCTAAACAGGT-----	229
Db	1015	TATGTTGGAACAAGGAGGAGATTCATTCCTGTTATCTCTAAACAGATCGCTCCAGAC	1074
Qy	230	-----	229
Db	1075	TCGCCCTACAGATAGTGTGATCTCACCTGAATTAATTTATCCAGTAGAAGATGATTTAT	1134
Qy	230	-----ACAGCAATCTTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGC	279
Db	1135	GGAAAGGCAGAGACTCATTTAGTCAGTGCCTGATTCGTGACGCCCACTAAATTAATCTATGA	1194
Qy	280	TGAAGGAGGATCAT-----AGTTGATGAAGTCAGAAAGCTGGTATATT	324
Db	1195	AGAAACAACAGACATCTGGATAAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATT	1254
Qy	325	TGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAGTTAGTAAATCC	384
Db	1255	TGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGCTAGTCAGTTAGTAAATCC	1314
Qy	385	TGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACTTCTTGTGTCATCAGTCAGCA	444
Db	1315	TGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACTTCTTGTGTCATCAGTCAGCA	1374
Qy	445	CTGTGACTCTTTTAAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTA	504
Db	1375	CTGTGACTCTTTTAAAGTAAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTA	1434
Qy	505	CAAGCTATCAAGTCTGAAGATGACCAACTTCGAAACAAAGGAATTTTGGGCCACCAT	564
Db	1435	CAAGCTATCAAGTCTGAAGATGACCAACTTCGAAACAAAGGAATTTTGGGCCACCAT	1494
Qy	565	TTTGGATTACAGAGTCTCTTCTGACTATATCTCCCTCAGAAATTTTCTTTTGAAG	624
Db	1495	TTTGGATTACAGAGTCTCTTCTGACTATATCTCCCTCAGAAATTTTCTTTTGAAG	1554
Qy	625	TACTACTGATTTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAA	684
Db	1555	TACTACTGATTTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAA	1614
Qy	685	GAAATATCTACTGTGCTTATATATGTTGGTGTCTCAGGTGCGAGTTGGTGAATAATCG	744
Db	1615	GAAATATCTACTGTGCTTATATATGTTGGTGTCTCAGGTGCGAGTTGGTGAATAATCG	1674
Qy	745	GTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGCTGTTGT	804
Db	1675	GTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGCTGTTGT	1734
Qy	805	AGTGATAGCAACAGGGATCCTGTGACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATA	864
Db	1735	AGTGATAGCAACAGGGATCCTGTGACCGAGGGCTTAAATTTGAAGGGCCCTTTAAATA	1794
Qy	865	TAAATAGGATCAATAGAAATTTGATGAGTGGAGGAGTCCATATCTAGTCTTCG	924
Db	1795	TAAATAGGATCAATAGAAATTTGATGAGTGGAGGAGTCCATATCTAGTCTTCG	1854
Qy	925	ATATGATTTCAATGACTTATGATCGTGTGGGATCCAGCGTGGTCTATGAGGAGTACCT	984
Db	1855	ATATGATTTCAATGACTTATGATCGTGTGGGATCCAGCGTGGTCTATGAGGAGTACCT	1914
Qy	985	CTCCCTGATGGCATTAATGACAGGCTCAGATATCTTCAGGGTGTCTTATTCCTGGGGCCCC	1044
Db	1915	CTCCCTGATGGCATTAATGACAGGCTCAGATATCTTCAGGGTGTCTTATTCCTGGGGCCCC	1974
Qy	1045	AGTCACTCTGTGGATCTTCTATGATACAGATACACGGA	1083
Db	1975	AGTCACTCTGTGGATCTTCTATGATACAGATACACGGA	2013
RESULT 6			
AX354795			
LOCUS	AX354795	2649 bp	DNA linear PAT 06-FEB-2002
DEFINITION	Sequence 3 from Patent WO0179473.		
ACCESSION	AX354795		
VERSION	AX354795.1	GI:18619528	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Meyers, R.A. and Williamson, M.		
TITLE	21953, a human prolyl oligopeptidase family member and uses thereof		
JOURNAL	Patent: WO 0179473-A 3 25-OCT-2001;		
	Millennium Pharmaceuticals, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..2649		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
	Query Match	73.1%; Score 791.2; DB 6; Length 2649;	
	Best Local Similarity	99.0%; Pred. No. 2.2e-184;	
	Matches	796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
Qy	280	TGAAGGAAGGATCATAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGA	339
Db	1527	TGGATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGA	1586
Qy	340	CTCCCTTTAGAGCATCACCTGTAGTACGTAGTAAATCCTGGAGAGTGACAAG	399
Db	1587	CTCCCTTTAGAGCATCACCTGTAGTACGTAGTAAATCCTGGAGAGTGACAAG	1646
Qy	400	GCTGACTCAGCGTGGCTACTCACATTTCTGTCATCAGTCAGCAGTCTGTTCTTTAT	459
Db	1647	GCTGACTCAGCGTGGCTACTCACATTTCTGTCATCAGTCAGCAGTCTGTTCTTTAT	1706
Qy	460	AAAGTAAATATAGTAACCAAGAAATCCACTGTGTGCTCCCTTTACAAGCTATCAAGTCC	519
Db	1707	AAAGTAAATATAGTAACCAAGAAATCCACTGTGTGCTCCCTTTACAAGCTATCAAGTCC	1766
Qy	520	TGAAGATGACCAACTTTGCAAAAACAAAGGAATTTTGGGCCACCATTTGGATTACAGAGG	579
Db	1767	TGAAGATGACCAACTTTGCAAAAACAAAGGAATTTTGGGCCACCATTTGGATTACAGAGG	1826
Qy	580	TCCTCTCTGATATATCTCTCAGAAATTTCTCTTTTGAAGTACTACTGGATTAC	639
Db	1827	TCCTCTCTGATATATCTCTCAGAAATTTCTCTTTTGAAGTACTACTGGATTAC	1886
Qy	640	ATTGATGGGATGCTCTTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTCT	699
Db	1887	ATTGATGGGATGCTCTTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTCT	1946
Qy	700	GCTGTTTATATATGTTGCTCTCAGGTGCTAGTGGTGAATAATCGGTTTAAAGAGTCAA	759
Db	1947	GCTGTTTATATATGTTGCTCTCAGGTGCTAGTGGTGAATAATCGGTTTAAAGAGTCAA	2006
Qy	760	GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTATAGTACACAG	819
Db	2007	GTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTATAGTACACAG	2066
Qy	820	GGGATCTCTGTCCAGGGCTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAAAT	879







```
Db 2389 CTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATT 2448
Qy 1000 AATCAGAGGTGAGATATCTTCAAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGTGAT 1059
Db 2449 AATCAGAGGTGAGATATCTTCAAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGTGAT 2508
Qy 1060 CTTCTATGATACAGGATACACGGA 1083
Db 2509 CTTCTATGATACAGGATACACGGA 2532

RESULT 10
AR651452
LOCUS
DEFINITION Sequence 2 from patent US 6881564.
ACCESSION AR651452
VERSION AR651452.1 GI:62795938
KEYWORDS
SOURCE Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 3120)
AUTHORS Abbott C.A. and Gorrell, M.D.
TITLE Dipeptidyl peptidases
JOURNAL Patent: US 6881564-A 2 19-APR-2005;
The University of Sydney; Sydney;
AUX;
FEATURES
source
1..3120
/organism="unknown"
/locus_tag="genomic DNA"
ORIGIN
Query Match 73.1%; Score 791.2; DB 6; Length 3120;
Best Local Similarity 99.0%; Pred. No. 2.2e-184;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 280 TGAAGGAAGCATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
Db 1740 TGAGTATAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1799
Qy 340 CTCCTCTTTAGAGCATCACTGTAGTACGTAGTACGTAAATCTTGAAGAGGTGACAG 399
Db 1800 CTCCTCTTTAGAGCATCACTGTAGTACGTAGTACGTAAATCTTGAAGAGGTGACAG 1859
Qy 400 GCTGACTGACCGTGGCTACTCAGATTCCTGTCATCAGTCAGTCAGTCAGTCAGTCAGTC 459
Db 1860 GCTGACTGACCGTGGCTACTCAGATTCCTGTCATCAGTCAGTCAGTCAGTCAGTCAGTC 1919
Qy 460 AAGTAAGTATAGTAAACCAAGAAATCCACACTGTGTCTCTTTTCAAGCTATCAAGTCC 519
Db 1920 AAGTAAGTATAGTAAACCAAGAAATCCACACTGTGTCTCTTTTCAAGCTATCAAGTCC 1979
Qy 520 TGAAGATGACCAACTCTCCAAACCAAGGAATTTGGGCCACCACTTTTGGATTCAGCAGG 579
Db 1980 TGAAGATGACCAACTCTCCAAACCAAGGAATTTGGGCCACCACTTTTGGATTCAGCAGG 2039
Qy 580 TCCTCTCTGACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTAC 639
Db 2040 TCCTCTCTGACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTAC 2099
Qy 640 ATTGTATGGATGCTCTCAAGCTCTATGATCTACAGCTCTGGAAGAAATATCTTACTGT 699
Db 2100 ATTGTATGGATGCTCTCAAGCTCTATGATCTACAGCTCTGGAAGAAATATCTTACTGT 2159
Qy 700 GCTGTTCTATATATGTTGCTCTCAGGTGCAAGTGGTGAATAATTCGGTTTAAAGAGTCAA 759
Db 2160 GCTGTTCTATATATGTTGCTCTCAGGTGCAAGTGGTGAATAATTCGGTTTAAAGAGTCAA 2219
Qy 760 GTATTTCCGCTGTAATACCCCTAGCTCTCTAGGTTATGTTGGTTAGTAGATAGACAAAG 819
Db 2220 GTATTTCCGCTGTAATACCCCTAGCTCTCTAGGTTATGTTGGTTAGTAGATAGACAAAG 2279

RESULT 11
AF221634
LOCUS
DEFINITION Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, complete cds.
ACCESSION AF221634
VERSION AF221634.1 GI:11095187
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
1 (bases 1 to 3127)
AUTHORS Abbott, C.A., Yu, D.M., Woollatt, E., Sutherland, G.R., McCaughan, G.W.
and Gorrell, M.D.
TITLE Cloning, expression and chromosomal localization of a novel human
dipeptidyl peptidase (DPP) IV homolog, DPP8
JOURNAL Biochem. 267 (20), 6140-6150 (2000)
REFERENCE
1 (bases 1 to 3127)
AUTHORS Abbott, C.A., Yu, D., McCaughan, G.W. and Gorrell, M.D.
JOURNAL Direct Submission
TITLE Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver
Centre, Centenary Institute of Cell Biology and Cancer Medicine,
Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia
FEATURES
source
1..3127
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q22"
1..3127
/gene="DPP8"
214..2862
/gene="DPP8"
/feature="DPP8"
/feature="serine protease; member of prolyl oligopeptidase S9b
family belonging to enzyme clan SC; alternatively spliced
forms of this mRNA are found in AF221635, AF221636 and
AF221637"
/codon_start=1
/product="dipeptidyl peptidase 8"
/protein_id="AAG29766.1"
/db_xref="GI:11095188"
/translation="MAAAMTEQLGVETADCEENIESQDRPKLEPFYVRSWSQ
LKLLADTRKYVGYMAKAPHDFMFKRNDPDGPHSDRIYLYLMSGRENITLYSEI
PKTNRAALVLMKWLPLDFOATLDYGMYSREELRERKIGTVGTASVDYHQSG
TFLQAGSLYHVKDGPGQFTOOPRLNIVETSCNIRMDPKIAPADDMATFTHSN
DIWLSNIVTREERLTYVHNELANMEEDARSAGVATFVLOREDFYSQYWCPRKAIT
PSGKILRLVYEENDESEVILHTVSPMLERADSFYKATNPKYTFKMSIIMI
DAEGRIIDVLKSLQPFLLFEGVYIARAGWTEGKYANSLIDRSQIRLIQVLLIS
PELFIPEDDVMERQRLIESVPSVTLIYEETDINIHDIHFVFPQSHEEIEF
```

ORIGIN									
Query Match 73.1%; Score 791.2; DB 8; Length 3127;									
Best Local Similarity 99.0%; Pred. No. 2.2e-184;									
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;									
QY	280	TGAAGGAGGATCATAGTTGATGAGTACAGAGGCTGGTATATTTTGAAGGCACCAAGA	339						
DB	1740	TGGATCTAATATCAAGTTGATGAGTACAGAGGCTGGTATATTTTGAAGGCACCAAGA	1799						
QY	340	CTCCCTTTAGAGCATCAGCTGTACGTAGTACGTATTAATCTTGAGAGGTGACAAG	399						
DB	1800	CTCCCTTTAGAGCATCAGCTGTACGTAGTACGTATTAATCTTGAGAGGTGACAAG	1859						
QY	400	GCTGACTGACCGTGGCTACTCACTTCTGTGCTAGTACGTACGACCTGTGACTTCTTTAT	459						
DB	1860	GCTGACTGACCGTGGCTACTCACTTCTGTGCTAGTACGTACGACCTGTGACTTCTTTAT	1919						
QY	460	AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCC	519						
DB	1920	AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCC	1979						
QY	520	TGAAGATGACCCAACTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTGACGAGG	579						
DB	1980	TGAAGATGACCCAACTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTGACGAGG	2039						
QY	580	TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAAAAGTACTACTGGATTTAC	639						
DB	2040	TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAAAAGTACTACTGGATTTAC	2099						
QY	640	ATTGTATGGAGTGTCTACAAGCCTCATGATCTACAGCTCGAAAGAAATATCTACTCTGT	699						
DB	2100	ATTGTATGGAGTGTCTACAAGCCTCATGATCTACAGCTCGAAAGAAATATCTACTCTGT	2159						
QY	700	GCTGTTTCATATATGGTGTCTCAGGTGCAGTGTGGTGAATAATCGGTTTAAAGGAGTCAA	759						
DB	2160	GCTGTTTCATATATGGTGTCTCAGGTGCAGTGTGGTGAATAATCGGTTTAAAGGAGTCAA	2219						
QY	760	GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGTTGATGATAGACAACAG	819						
DB	2220	GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGTTGATGATAGACAACAG	2279						
QY	820	GGGATCTCTGTACACGAGGGCTTAAATTTGAAGGCGCCTTTAAATATAAAATGGGTCAAAT	879						
DB	2280	GGGATCTCTGTACACGAGGGCTTAAATTTGAAGGCGCCTTTAAATATAAAATGGGTCAAAT	2339						
QY	880	AGAAATTCACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGA	939						
DB	2340	AGAAATTCACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGA	2399						
QY	940	CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATT	999						
DB	2400	CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATT	2459						
QY	1000	AATGACAGGTCAGATATCTTACAGGTTGCTATGCTGGGGCCCCAGTCACTCTGTGGAT	1059						
DB	2460	AATGACAGGTCAGATATCTTACAGGTTGCTATGCTGGGGCCCCAGTCACTCTGTGGAT	2519						
QY	1060	CTTCTATGATACAGGATACACGGA 1083							
DB	2520	CTTCTATGATACAGGATACACGGA 2543							
RESULT 12									
AX354793									

LOCUS	AX354793	3143 bp	DNA	linear	PAT 06-FEB-2002
DEFINITION	Sequence 1 from Patent WO0179473.				
ACCESSION	AX354793				
VERSION	AX354793.1	GI:18619526			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1				
AUTHORS	Meyers, R.A. and Williamson, M.				
TITLE	21953, a human prolyl oligopeptidase family member and uses thereof				
JOURNAL	Patent: WO 0179473-A 1 25-OCT-2001;				
Millennium Pharmaceuticals, Inc. (US)					
FEATURES	Location/Qualifiers				
source	1..3143				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
CDS	229..2877				
	/notes="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD22990.1"				
	/db_xref="GI:18619527"				
	/translation="MAAAWETBQLGVEIFETADCEENIESQDRPKLPFFYVERYSWQ LKLLADTRKYHGMKAPLDFMFKRNDPDGSHSDRIYYLAMSNGNRENTLFYSEI PKTINRAAVLMSWPKPLDLFOATLDYGVYSREELLERKRIGTVGIASDYHOGSG TFIFOAGSGIYHVKGQPGQFTQOPLPNLVETSPNMDPKLPADPWIAFHNS DWINIVTREERRLLTYVHNELANNEEDARSAGVATFVLOEFDRYSGYMWCPKAEIT PSGKILRLIYBENSEVEIILHVSPLMETRRADSFRYPKGTANTPKVTFKMEIMI DAGRIIDVIDKELIQPFELFEGVEYIARAGWTPEGKYAWSILLDRSQTKRLQIVLIS PELFIPEDDVMERQRLIESVPSVTPLIIEETDIWINIHDIHFVFPQSHSEELF IFASECKTGRHLHYKITSILKESKYKRSGLPAPSPDFKCPKEEIAITSGWEVLGR HGSNIQDVSRVLVVEGKDSPLHHLVYVYVNPGEVTRLDRGVSHSCCISOHCD PFLTSKSNQKPHCVSLYKLSPEDEPDKTKKFWAILDSAGPLDYPPELSPES TTGFTLLYGLMYKPHDLQPKKYPTVLFIYGGPQVQLVNNRFGVKYFRNLTLASLGYV VVVIDNRGSGHRLKPEGAFFKMGQIBIDDOVEGLQYLASRYDFDLDRVGHGWSY GGVLSLMALMORSIDIFRVAIAGAPVTLMIFYDTGYTRYMGHPDQEQGYLGSVAMQ AEKFPSEPNRLLLHGFIDENVFAHTSILLSFLVRAGKPYDLQIYQPERHSIRVPES GEHYELHLLHYLQENLGSRIAALKVI"				
ORIGIN					
Query Match	73.1%; Score 791.2; DB 6; Length 3143;				
Best Local Similarity	99.0%; Pred. No. 2.2e-184;				
Matches	796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;				
Qy	280	TGAAGGAGGATCATAGTTGATGAGTACAGAGGCTGGTATATTTTGAAGGCACCAAGA	339		
Db	1755	TGGATCTAATATCCAGTTGATGAGTACAGAGGCTGGTATATTTTGAAGGCACCAAGA	1814		
Qy	340	CTCCCTTTAGAGCATCAGCTGTACGTAGTACGTATTAATCTTGAGAGGTGACAAG	399		
Db	1815	CTCCCTTTAGAGCATCAGCTGTACGTAGTACGTATTAATCTTGAGAGGTGACAAG	1874		
Qy	400	GCTGACTGACCGTGGCTACTCACTTCTGCTGCTAGTACGTAAATCTCGAGAGGTGACAAG	459		
Db	1875	GCTGACTGACCGTGGCTACTCACTTCTGCTGCTAGTACGTAAATCTCGAGAGGTGACAAG	1934		
Qy	460	AAGTAAGTATAGTAACCAAGAAATTTTGGGCCACCAATTTTGGATTGACGAGG	519		
Db	1935	AAGTAAGTATAGTAACCAAGAAATTTTGGGCCACCAATTTTGGATTGACGAGG	1994		
Qy	520	TGAAGATGACCAACTTTGCAAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTGACGAGG	579		
Db	1995	TGAAGATGACCAACTTTGCAAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTGACGAGG	2054		
Qy	580	TCCTCTTCTGACTATACTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGATTAC	639		
Db	2055	TCCTCTTCTGACTATACTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGATTAC	2114		
Qy	640	ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGT	699		

Db 2115 ATGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGGAAGAATAATCTCTACTGT 2174

Qy 700 GCTGTTTCATATATGTTGGTCTCTCAGGTGCGATGTTGGTGAATAATCGGTTTAAAGAGTCAA 759

Db 2175 GCTGTTTCATATATGTTGGTCTCTCAGGTGCGATGTTGGTGAATAATCGGTTTAAAGAGTCAA 2234

Qy 760 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGGTGTAGTAGTAGACAACAG 819

Db 2235 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGGTGTAGTAGTAGACAACAG 2294

Qy 820 GGGATCCTGTACCCAGGGCTTAAATTTGAAGCGCTTAAATATATAAATGGTCAAT 879

Db 2295 GGGATCCTGTACCCAGGGCTTAAATTTGAAGCGCTTAAATATATAAATGGTCAAT 2354

Qy 880 AGAATTTGACGATCAGTGAAGGACTCCCAATATCTAGCTTCTGATGATGATTCATGA 939

Db 2355 AGAATTTGACGATCAGTGAAGGACTCCCAATATCTAGCTTCTGATGATGATTCATGA 2414

Qy 940 CTTAGATCGTGGGATCCACCGCTGCTCTATGAGGATACCTCTCCCTGATGGATT 999

Db 2415 CTTAGATCGTGGGATCCACCGCTGCTCTATGAGGATACCTCTCCCTGATGGATT 2474

Qy 1000 AATCAGAGTCAAGATATCTTCAGGTTCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 1059

Db 2475 AATCAGAGTCAAGATATCTTCAGGTTCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 2534

Qy 1060 CTTCTATGATACAGGATACACGGA 1083

Db 2535 CTTCTATGATACAGGATACACGGA 2558

RESULT 13

BC040203

LOCUS Homo sapiens dipeptidylpeptidase 8, mRNA (cDNA clone MGC:42570 IMAGE:4824813), complete cds.

DEFINITION

ACCESSION BC040203

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 4535)

Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Worrley, K.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalios, D.E., Schnarch, A., Schein, J.E., Jones, S.J. and Warra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 4535)

Strausberg, R.

Direct Submission

Submitted (27-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK

COMMENT

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbe-r@mail.nih.gov](mailto:cgapbe-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (NIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amgebcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleeged, H., Kowis, C.R., Sreed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 71 Row: k Column: 23.

Location/Qualifiers

FEATURES

source

1..4535

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:42570 IMAGE:4824813"

/tissue\_type="Testis"

/clone\_lib="NIH MGC\_97"

/lab\_host="DH10B"

/notes="Vector: pBluescript"

gene

1..4535

/gene="DPP8"

/notes="synonyms: FLJ20283, FLJ14920, MGC26191"

/db\_xref="GeneID:54878"

/db\_xref="MIM:606819"

1623..4277

/gene="DPP8"

/codon\_start=1

/product="DPP8 protein"

/protein\_id="AAH40203.1"

/db\_xref="GI:26007917"

/db\_xref="GeneID:54878"

/db\_xref="MIM:606819"

/translation="MAAAMETEQLGVEIFADCEENIESQDRPKLPFFVYRWSQ LKLLADTRKHYGMWAKAPHDMEVKNDPDGPHSDRIYLYLMMRENTLIFYSEI PTKINRAAVLMLSWKPLLDLQATLDYMGYSREELLRKRKIGTVGIAGSYHQSSG TLFQAGSGIYHVKGQGGTQQLPRLNVLVETSCPNRMDPKLCPADPDWIAFHGN DWISNIVTREERLTYVHNELANMEEDARSAGVATFVLOEFDRYSGYVWPKAEIT PSQKTLIRLYEENDESEIIVHTSPMLERADSPRYKPTGTANPKVTFKMSIEMI DAEGRIDVIDKELIQPEILFEGVEYLARAGTPEGYKWSILLDRSQRLQIVLIS PELFIVEDVMERQRLIESVPSVTPLIYEETDIIWINHDIHFVFPQSHBEETEF IFASEKTFGRHLXKITSLKSKYKSSGGLPAPSDPKCPKEEIAITSGEWELGR HGSNIQDVRRLVVFEGTKDSPLEHLYVYVYVNGEVTRLTDRGYSHCISQHCD FTISKYKNKNPHCVSLYKLSSPDDPTCKTFEATILDSAGPLPDYTPPELSPFES TTGFTLYGLYKPHLOPKKYPTVLFIYGGPOVLNRPKGVKVPRLATLASGVY VVVIDNRGSHRGKFEAGFYKKGKGLIEDDQVGLQYLSRVDFDLDVRVTHGVY GGLYSIMALMORSDIFRVAIAGA PVTLMWIFVDTGYTRYMGHPDQNEQGYLGSVAMQ AKKFPSEPNRLLLHGFUDENVFAHTSILLSFLVRAGKPYDLQIIPQERHSIRVPE GEHYELHLHLVLOENLGSRIAALKVI"

CDS

1623..4277

/gene="DPP8"

/codon\_start=1

/product="DPP8 protein"

/protein\_id="AAH40203.1"

/db\_xref="GI:26007917"

/db\_xref="GeneID:54878"

/db\_xref="MIM:606819"

/translation="MAAAMETEQLGVEIFADCEENIESQDRPKLPFPFVYRWSQ LKLLADTRKHYGMWAKAPHDMEVKNDPDGPHSDRIYLYLMMRENTLIFYSEI PTKINRAAVLMLSWKPLLDLQATLDYMGYSREELLRKRKIGTVGIAGSYHQSSG TLFQAGSGIYHVKGQGGTQQLPRLNVLVETSCPNRMDPKLCPADPDWIAFHGN DWISNIVTREERLTYVHNELANMEEDARSAGVATFVLOEFDRYSGYVWPKAEIT PSQKTLIRLYEENDESEIIVHTSPMLERADSPRYKPTGTANPKVTFKMSIEMI DAEGRIDVIDKELIQPEILFEGVEYLARAGTPEGYKWSILLDRSQRLQIVLIS PELFIVEDVMERQRLIESVPSVTPLIYEETDIIWINHDIHFVFPQSHBEETEF IFASEKTFGRHLXKITSLKSKYKSSGGLPAPSDPKCPKEEIAITSGEWELGR HGSNIQDVRRLVVFEGTKDSPLEHLYVYVYVNGEVTRLTDRGYSHCISQHCD FTISKYKNKNPHCVSLYKLSSPDDPTCKTFEATILDSAGPLPDYTPPELSPFES TTGFTLYGLYKPHLOPKKYPTVLFIYGGPOVLNRPKGVKVPRLATLASGVY VVVIDNRGSHRGKFEAGFYKKGKGLIEDDQVGLQYLSRVDFDLDVRVTHGVY GGLYSIMALMORSDIFRVAIAGA PVTLMWIFVDTGYTRYMGHPDQNEQGYLGSVAMQ AKKFPSEPNRLLLHGFUDENVFAHTSILLSFLVRAGKPYDLQIIPQERHSIRVPE GEHYELHLHLVLOENLGSRIAALKVI"

misc\_feature

3207..4196

/gene="DPP8"

/note="DAP2; Region: Dipeptidyl aminopeptidases/acylaminoacyl-peptidases [Amino acid transport and metabolism]"

/db\_xref="CDD:COG1506"

ORIGIN

Query Match 73.1%; Score 791.2; DB 8; Length 4535;

Best Local Similarity 99.0%; Pred. No. 2.1e-184;

Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 280 TGAAGGAGGATCATAGTTGATGAAGTCAAGGCTGGTATATTTTGAAGGCCCAAGA 339

|||||

3155 TGAGTCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGA 3214  
Db  
340 CTCCTCTTTAGAGCATCACCTGTAGTACGTACGTAAATCTCGAGAGGTGACAA 399  
Qy  
3215 CTCCTCTTTAGAGCATCACCTGTAGTACGTAAATCTCGAGAGGTGACAA 3274  
Db  
400 GCTGACTGACCGTGGCTACTACATCTTCTGCTGATCAGTCAGCATGCTGCTTTAT 459  
Qy  
3275 GCTGACTGACCGTGGCTACTACATCTTCTGCTGATCAGTCAGCATGCTGCTTTAT 3334  
Db  
460 AGTAAATGATATAGTACCAAGAAATCCACATCTGTGCTCCCTTTTCAAGCTATCAAGTCC 519  
Qy  
3335 AAGTAAATGATATAGTACCAAGAAATCCACATCTGTGCTCCCTTTTCAAGCTATCAAGTCC 3394  
Db  
520 TGAAGATGACCCAACTTCCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGG 579  
Qy  
3395 TGAAGATGACCCAACTTCCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGG 3454  
Db  
580 TCTCTCTGCTGATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTAC 639  
Qy  
3455 TCTCTCTGCTGATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTAC 3514  
Db  
640 ATTGATGCGGATGCTCTCAAGCTCTATGATCTACAGCTGGAAAGAAATATCTACTGT 699  
Qy  
3515 ATTGATGCGGATGCTCTCAAGCTCTATGATCTACAGCTGGAAAGAAATATCTACTGT 3574  
Db  
700 GCTGTTTCATATATGTTGGTCTCAGGTCAGTGGTGAATATCGGTTTAAAGGAGTCAA 759  
Qy  
3575 GCTGTTTCATATATGTTGGTCTCAGGTCAGTGGTGAATATCGGTTTAAAGGAGTCAA 3634  
Db  
760 GTATTTCCGTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTATGATACCAACAG 819  
Qy  
3635 GTATTTCCGTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTATGATACCAACAG 3694  
Db  
820 GGCATCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAAT 879  
Qy  
3695 GGCATCTGTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAAT 3754  
Db  
880 AGAAATGACGATCAGTGGAGGACTCCAAATATAGCTTCTCGATATGATTTCAATTGA 939  
Qy  
3755 AGAAATGACGATCAGTGGAGGACTCCAAATATAGCTTCTCGATATGATTTCAATTGA 3814  
Db  
940 CTAGATCTGTTGGGATCCAGGCTGCTCTAGGAGGATACCTCTCCCTGATGGCATT 999  
Qy  
3815 CTAGATCTGTTGGGATCCAGGCTGCTCTAGGAGGATACCTCTCCCTGATGGCATT 3874  
Db  
1000 AATCGAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCATCTGTGGAT 1059  
Qy  
3875 AATCGAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCGAGTCATCTGTGGAT 3934  
Db  
1060 CTTCTATGATACAGGATACACGA 1083  
Qy  
3935 CTTCTATGATACAGGATACACGA 3958  
Db  
RESULT 14  
AV172659 2656 bp mRNA linear PRI 08-JAN-2003  
LOCUS  
DEFINITION Homo sapiens dipeptidyl peptidase IV-related protein-1 (DPP1)  
mRNA, complete cds.  
ACCESSION  
VERSION AV172659.1 GI:27549549  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 2656)  
Qi, S., Akinsanya, K., Riviere, P. and Junien, J.-L.  
Novel Serine Protease Genes Related To DPP1V  
Patent: US (WO 0231134)-A 18-APR-2002;  
2 (bases 1 to 2656)

AUTHORS Qi, S., Akinsanya, K., Riviere, P. and Junien, J.-L.  
Direct Submission  
Submitted (04-NOV-2002) Ferring Research Institute, 3550 General  
JOURNAL Atomics Ct., San Diego, CA 92121, USA  
FEATURES  
source  
1..2656  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chTosome="15"  
/map="15q22.1-q22.2"  
1..2656  
/gene="DPP1"  
8..2656  
/gene="DPP1"  
/codon\_start=1  
/product="dipeptidyl peptidase IV-related protein-1"  
/protein\_id="AA017261.1"  
/translation="MAAMETEQLGVIFETACBENISQDRPKLPFFVVERYSWQ  
LKKLADTKYHGMKAKPHDFMFKRNDPDGSDRIYVLTAMSGENRENTLFSYSEI  
PKTINRAVLMSKPLDLFOATLDYGMYSREELLRRKRIGTIGIASYDHQSG  
TFLFAGSGIYHVDGPGQFTQOPLSPNIVETSCPNRMDPKLCPADWDPAFTHSN  
DIWISNIVREERLLTVHNELANMEDARSAGVATFVLQEDRFYKGTANPKYTFKMSIMI  
DAEGRIIDVIDKELIQPEILFEFGEYIARAGWTPGKYAWMSILLDRSQTRQLVLIS  
PELFTPDDVMERQRLIESVPDSVTPLLIYEETDIWINDIHFVFSHEEELF  
IFASCKTGFRHLVITSLKSKYKRSGLPAPDFKCPKIEELATSGBEVLGR  
HGSNTQDVEVRLLVYFSGTQDPLHLYVYVNGEVTRLDRGYSHSCCTISOHCD  
FPIKSNQKPHCVSLYKLSBDEDDTCTKSFATILDSAGPLDPTPPEIFSPES  
VTGFTLYGKLPKHPDQKPKYPTVLVFIGGQVQLVNNRPKGVKFRNLTLASLVY  
VVVINDRSGHRLKDFQAGKFGYKMOIEIDQVEGLQYLASRYDFIDLRVGHGHSY  
AGKLSMALMORSIDIFRVAIAGAPVTLWIFDYTERYMGHPDQNEQYGLGSAVQ  
AKFSEPNRLILLHPLDENVHFANTSIILLFLVRAGKPYDLQIYPOEHSIRVPES  
GEHYELHLLHQLQENLGSRIAAKVI"

## ORIGIN

Query Match 72.9%; Score 789.6; DB 8; Length 2656;  
Best Local Similarity 98.9%; Pred. No. 5.4e-184;  
Matches 795; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
Qy 280 TGAAGAGAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGA 339  
Db 1534 TGGATCTAATATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGA 1593  
Qy 340 CTCCTCTTAGAGCATCACCTGTAGCTAGTACCTTACCTAAATCTCGAGAGGTGACAAG 399  
Db 1594 CTCCTCTTAGAGCATCACCTGTAGCTAGTACCTTACCTAAATCTCGAGAGGTGACAAG 1653  
Qy 400 GCTGACTGACCGTGGCTACTCACATTTCTTGTGCATCATGTCAGCTGTGACTCTTTAT 459  
Db 1654 GCTGACTGACCGTGGCTACTCACATTTCTTGTGCATCATGTCAGCTGTGACTCTTTAT 1713  
Qy 460 AAGTAAGTATAGTAACAGAGNAATCCACACTGTGCTCCCTTTTACAAGCTATCAAGTCC 519  
Db 1714 AAGTAAGTATAGTAACAGAGNAATCCACACTGTGCTCCCTTTTACAAGCTATCAAGTCC 1773  
Qy 520 TGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCATTTTGGATTCAGCAGG 579  
Db 1774 TGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCATTTTGGATTCAGCAGG 1833  
Qy 580 TCCTCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTAC 639  
Db 1834 TCCTCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTAC 1893  
Qy 640 ATTGTATGGGATGCTCTTACAGGCTCATGATCTACAGCTGGAAAGAAATATCTTACTGT 699  
Db 1894 ATTGTATGGGATGCTCTTACAGGCTCATGATCTACAGCTGGAAAGAAATATCTTACTGT 1953  
Qy 700 GCTGTTTCATATATGTTGGTCTCCTCAGGTCAGTGGTGAATAATCGGTTTAAAGGAGTCAA 759  
Db 1954 GCTGTTTCATATATGTTGGTCTCCTCAGGTCAGTGGTGAATAATCGGTTTAAAGGAGTCAA 2013

QY 760 GTATTTCGGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAG 819  
Db 2014 GTATTTCGGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAG 2073  
QY 820 GGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCCTTAAATATAAATGAGTCAAT 879  
Db 2074 GGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCCTTAAATATAAATGAGTCAAT 2133  
QY 880 AGAAATTGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 939  
Db 2134 AGAAATTGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 2193  
QY 940 CTTAGATCGTGTGGGATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAAT 999  
Db 2194 CTTAGATCGTGTGGGATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAAT 2253  
QY 1000 AATGACAGGTGAGATATCTTCAAGGTTGCTATTCCTGGGGCCCAAGTCACTCTGTGGAT 1059  
Db 2254 AATGACAGGTGAGATATCTTCAAGGTTGCTATTCCTGGGGCCCAAGTCACTCTGTGGAT 2313  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
Db 2314 CTTCTATGATACAGGATACACGGA 2337

RESULT 15  
AR631278  
LOCUS AR631278 2671 bp DNA linear PAT 14-FEB-2005  
DEFINITION Sequence 2 from patent US 6844180.  
ACCESSION AR631278  
VERSION AR631278.1 GI:59770920  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2671)  
AUTHORS OI,S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.  
TITLE Serine protease genes related to DPPiV  
JOURNAL Patent: US 6844180-A 2 18-JAN-2005;  
Ferring BV;;  
NLX;

FEATURES  
source  
1..2671  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 72.9%; Score 789.6; DB 6; Length 2671;  
Best Local Similarity 98.9%; Pred. No. 5.4e-184;  
Matches 795; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 280 TGAAGGAAGGATCATAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCCACCAAGA 339  
Db 1534 TGGATCTAATATCCNAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCCACCAAGA 1593  
QY 340 CTCCTCTTTAGAGATCACTGTACGTAGTCAGTTACGTTAAATTCCTGGAGAGGTGACAAG 399  
Db 1594 CTCCTCTTTAGAGATCACTGTACGTAGTCAGTTACGTTAAATTCCTGGAGAGGTGACAAG 1653  
QY 400 GCTGACTGACCGTGGCTACTCACAATCTTCTGCTCATCAGTCAGCACTGTGACTTCTTTAT 459  
Db 1654 GCTGACTGACCGTGGCTACTCACAATCTTCTGCTCATCAGTCAGCACTGTGACTTCTTTAT 1713  
QY 460 AAGTAAGPATAGTAACAGAGAATCCACACTGTGTGCTCCCTTTACAAAGCTATCAAGTCC 519  
Db 1714 AAGTAAGPATAGTAACAGAGAATCCACACTGTGTGCTCCCTTTACAAAGCTATCAAGTCC 1773  
QY 520 TGAAGATGACCCAACTTGGCAAAACAAAGGAATTTGGGGCCACCATTTTGGATTGAGCAGG 579  
Db 1774 TGAAGATGACCCAACTTGGCAAAACAAAGGAATTTGGGGCCACCATTTTGGATTGAGCAGG 1833  
QY 580 TCCTCTTCCTGACTATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTAC 639

Db 1834 TCCTCTTCCTGACTATATCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTAC 1893  
QY 640 ATTGTATGGGATGCTCTACAAAGCCTCATGATCTACAGCCTGGAAGAAAATATCCTACTGT 699  
Db 1894 ATTGTATGGGATGCTCTACAAAGCCTCATGATCTACAGCCTGGAAGAAAATATCCTACTGT 1953  
QY 700 GCTCTTCATATATGGTGGTCTCAGGTGAGTTGGTGAATAATCCGTTTAAAGGAGTCAA 759  
Db 1954 GCTCTTCATATATGGTGGTCTCAGGTGAGTTGGTGAATAATCCGTTTAAAGGAGTCAA 2013  
QY 760 GTATTTCCGCTTGAATAACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAG 819  
Db 2014 GTATTTCCGCTTGAATAACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAG 2073  
QY 820 GGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCCTTAAATATAAATGAGTCAAT 879  
Db 2074 GGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCCTTAAATATAAATGAGTCAAT 2133  
QY 880 AGAAATTGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 939  
Db 2134 AGAAATTGACGATCAGGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 2193  
QY 940 CTTAGATCGTGTGGGATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAAT 999  
Db 2194 CTTAGATCGTGTGGGATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAAT 2253  
QY 1000 AATGACAGGTGAGATATCTTCAAGGTTGCTATTCCTGGGGCCCAAGTCACTCTGTGGAT 1059  
Db 2254 AATGACAGGTGAGATATCTTCAAGGTTGCTATTCCTGGGGCCCAAGTCACTCTGTGGAT 2313  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
Db 2314 CTTCTATGATACAGGATACACGGA 2337

Search completed: May 4, 2006, 18:52:12  
Job time : 5400 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:50:44 ; Search time 609.925 Seconds  
(without alignments)

11834.025 Million cell updates/sec

Title: US-10-825-632-8

Perfect score: 1083

Sequence: 1 ggaagaagatgccagatcag.....tatgatacaggatacacgga 1083

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

1: geneseqn1980s.\*

2: geneseqn1990s.\*

3: geneseqn2000s.\*

4: geneseqn2001as.\*

5: geneseqn2001bs.\*

6: geneseqn2002as.\*

7: geneseqn2002bs.\*

8: geneseqn2003as.\*

9: geneseqn2003bs.\*

10: geneseqn2003cs.\*

11: geneseqn2003ds.\*

12: geneseqn2004as.\*

13: geneseqn2004bs.\*

14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1083	100.0	1083	4 AAC85697	Aac85697 Nucleotid
2	1055.8	97.5	4309	6 ABK83328	Abk83328 cDNA enco
3	791.2	73.1	1821	6 ABV76411	Abv76411 Dipeptidy
4	791.2	73.1	2083	7 ADR41222	Adr41222 Human CD-
5	791.2	73.1	2643	6 AAH99935	Aah99935 Coding se
6	791.2	73.1	2649	8 ABX12255	Abx12255 cDNA enco
7	791.2	73.1	2649	14 ADV43981	Adv43981 Human psy
8	791.2	73.1	2797	12 ADL13374	Adl13374 Human ste
9	791.2	73.1	2929	10 ACA92425	Aca92425 DNA enco
10	791.2	73.1	3106	6 ABK12892	Abk12892 Human pro
11	791.2	73.1	3120	4 AAC85694	Aac85694 Nucleotid
12	791.2	73.1	3120	6 AAD38956	Aad38956 Human dip
13	791.2	73.1	3143	6 AAH99934	Aah99934 cDNA enco
14	789.6	72.9	2671	6 ABK83322	Abk83322 cDNA enco
15	779.2	71.9	4829	6 ABK83327	Abk83327 cDNA enco
16	608.8	56.2	2251	10 ADE79035	Ade79035 Human pro
17	581.8	53.7	2463	10 ACA92424	Aca92424 DNA enco
18	581.8	53.7	2510	6 AAD23843	Aad23843 Human pro
19	581.2	53.7	2952	10 ACA92421	Aca92421 DNA enco

20	579.6	53.5	2161	4 AAH15009	Aah15009 Human cDN
21	579.2	53.5	2696	12 ADI16386	Adi16386 Human pro
22	497.2	45.9	4685	6 ABK83332	Abk83332 cDNA enco
23	479.2	44.2	4676	6 ABK83331	Abk83331 cDNA enco
24	459	42.4	1197	4 AAC85695	Aac85695 Nucleotid
25	434.2	40.1	1669	4 AAC85696	Aac85696 Nucleotid
26	434.2	40.1	2349	14 ADV43982	Adv43982 Human psy
27	434.2	40.1	2668	6 ABN59775	Abn59775 Novel hum
28	434.2	40.1	2842	6 ABN59774	Abn59774 Novel hum
29	428.2	39.5	4523	6 ABK83325	Abk83325 cDNA enco
30	408.6	37.7	2702	13 ADT04072	Adt04072 Human pro
31	365.4	33.7	502	9 ACH32373	Ach32373 Human end
32	345.8	31.9	2952	6 ABK85090	Abk85090 DNA enco
33	345	31.9	1856	10 AD607574	Ad607574 Novel cod
34	344.2	31.8	2495	6 AAD38957	Aad38957 Human dip
35	344.2	31.8	2617	6 ABK83323	Abk83323 cDNA enco
36	344.2	31.8	2751	6 AAD38311	Aad38311 Murine di
37	344.2	31.8	2766	7 ADR41345	Adr41345 Human CD-
38	344.2	31.8	3024	6 AAD38954	Aad38954 Human dip
39	344.2	31.8	3047	6 ABK69113	Abk69113 DNA enco
40	344.2	31.8	3287	6 AAD38955	Aad38955 Alternati
41	344.2	31.8	3716	6 ABO75955	Abq75955 Human PM
42	344.2	31.8	4147	12 AD086753	Ado86753 Human tum
43	344.2	31.8	4147	13 AD84273	Ad84273 Human tum
44	344.2	31.8	4147	13 ACN40423	Acn40423 Tumour-as
45	344.2	31.8	4219	6 ABK83335	Abk83335 cDNA enco

ALIGNMENTS

RESULT 1

AAC85697

ID AAC85697 standard; cDNA; 1083 BP.

XX AAC85697;

XX 29-JUN-2001 (first entry)

DT Nucleotide sequence encoding human DPP8 244Glu-341Ile+515Val-776Thr.

DE Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;

XX dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;

KW growth hormone deficiency; glucose level; mucosal regeneration;

KW non-insulin dependent diabetes mellitus; glucose intolerance;

KW immunosuppression; ss.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 2..1081

FT /\*tag= a

FT /product= "Human DPP8 244Glu-341Ile+515Val-776Thr"

FT /note= "No stop codon given"

FT WO200119866-A1.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-AU001085.

XX 10-SEP-1999; 99AU-00002762.

PR 18-FEB-2000; 2000AU-00005709.

XX (UNSY ) UNIV SYDNEY.

PA Abbott CA, Gorell MD;

PI WPI; 2001-281520/29.

XX P-PSDB; ABA47190.

DR New human dipeptidyl aminopeptidase (DPP8) useful for cleaving

XX substrates, identifying inhibitors of DPP8 catalytic activity which have

PT



therapeutic uses, and for detecting activated T cells.

Claim 21; Page 75-76; 78pp; English.

The sequences given in AAC85695-97 encode fragments of human dipeptidyl aminopeptidase (DPP8). DPP8 has substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA. Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase, because it is capable of hydrolysing the peptide bond C-terminal to Pro in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is useful for cleaving a substrate, and for detecting an activated T cell which involves measuring the level of DPP8 gene expression in a T cell. The level of DPP8 expression is detected by detecting the amount of DPP8 RNA in the cell. It is also useful for identifying a molecule capable of inhibiting the cleavage of the substrate by DPP8. Molecules identified as inhibiting DPP8 catalytic activity may be useful for treating diarrhoea, growth hormone deficiency, lowering glucose levels in non-insulin dependent diabetes mellitus and other disorders involving glucose intolerance, enhancing mucosal regeneration and as immunosuppressants

Sequence 1083 BP; 312 A; 213 C; 247 G; 311 T; 0 U; 0 Other;

Query Match 100.0%; Score 1083; DB 4; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAAGAAATTTGATAG	60
DB	1	GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAAGAAATTTGATAG	60
QY	61	ATATTCCTGCTATTGGTGTGCTCCAAAAGCTGAAACAACTCCCAAGTGGTGTAATAATCT	120
DB	61	ATATTCCTGCTATTGGTGTGCTCCAAAAGCTGAAACAACTCCCAAGTGGTGTAATAATCT	120
QY	121	TAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCC	180
DB	121	TAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCC	180
QY	181	TATGCTGGAAACAGGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTACAGCAATCC	240
DB	181	TATGCTGGAAACAGGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTACAGCAATCC	240
QY	241	TAAAGTCACATTTTAAAGATGTCAGAAATATGATGATGCTGAAGGAGGATCATAGTTGA	300
DB	241	TAAAGTCACATTTTAAAGATGTCAGAAATATGATGATGCTGAAGGAGGATCATAGTTGA	300
QY	301	TGAAGTCAGAGGCTGCTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACT	360
DB	301	TGAAGTCAGAGGCTGCTATATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACT	360
QY	361	GTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACTGACCTGGCTACTC	420
DB	361	GTACGTAGTCAGTTACGTAATCTCGAGAGGTGACAGGCTGACTGACCTGGCTACTC	420
QY	421	ACATTCCTGCTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAA	480
DB	421	ACATTCCTGCTGCATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAA	480
QY	481	GAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAGATGACCCAACTGCAA	540
DB	481	GAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTGAAGATGACCCAACTGCAA	540
QY	541	AACAAAGAAATTTTGGGCCACCAATTTGGATTGAGGCTCCTCTCTGACTATAGTATCC	600
DB	541	AACAAAGAAATTTTGGGCCACCAATTTGGATTGAGGCTCCTCTCTGACTATAGTATCC	600
QY	601	TCCAGAAATTTCTCTTTTCAAAGTACTACTGATTTACATGTTATGGATGCTCTACAA	660
DB	601	TCCAGAAATTTCTCTTTTCAAAGTACTACTGATTTACATGTTATGGATGCTCTACAA	660
QY	661	GCCTCATGATCTACAGGCTCGAAAGAAATATCTACTGCTGTTTCATATATGTTGCTCC	720
DB	661	GCCTCATGATCTACAGGCTCGAAAGAAATATCTACTGCTGTTTCATATATGTTGCTCC	720

QY	721	TCAGGTCCAGTTCGGTGAATAATCGGTTTAAAGAGTCAAGTATTTTCGCTTGAATACCCCT	780
DB	721	TCAGGTCCAGTTCGGTGAATAATCGGTTTAAAGAGTCAAGTATTTTCGCTTGAATACCCCT	780
QY	781	AGCCTCTCTAGGTTATGTGGTTGTAGTAGACAAACAGGGGATCCTGTCCCGAGGGCT	840
DB	781	AGCCTCTCTAGGTTATGTGGTTGTAGTAGACAAACAGGGGATCCTGTCCCGAGGGCT	840
QY	841	TAAATTTGAAGGGCCCTTTTAAATATATAAATGGGTCAAAATAGAAATTCAGATCAGGTGGA	900
DB	841	TAAATTTGAAGGGCCCTTTTAAATATATAAATGGGTCAAAATAGAAATTCAGATCAGGTGGA	900
QY	901	AGGACTCCAATATCTAGCTTCTCGATATGATGATTTCACTGACTTAGATCTGTGGGCATCCA	960
DB	901	AGGACTCCAATATCTAGCTTCTCGATATGATGATTTCACTGACTTAGATCTGTGGGCATCCA	960
QY	961	CGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTTAATGCAGAGGTGAGATATCTT	1020
DB	961	CGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTTAATGCAGAGGTGAGATATCTT	1020
QY	1021	CAGGCTTCTTATGGTGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACAC	1080
DB	1021	CAGGCTTCTTATGGTGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACAC	1080
QY	1081	GGA 1083	
DB	1081	GGA 1083	
RESULT 2			
ABK83328			
ID	ABK83328	standard; cDNA; 4309 BP.	
XX	AC	ABK83328;	
XX	DT	12-AUG-2002 (first entry)	
XX	DE	cDNA encoding human DPP8-1 splice variant #4.	
XX	KW	Human; serine protease; dipeptidyl peptidase IV-related protein; DPP8;	
XX	KW	DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;	
XX	KW	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;	
XX	KW	heart failure; hypertension; urinary retention; osteoporosis; cancer;	
XX	KW	ulcer; allergy; cancer; psychotic disorder; neurological disorder;	
XX	KW	dyskinesia; reproductive disorder; inflammatory disorder;	
XX	KW	metabolic disorder; gene; ss.	
OS		Homo sapiens.	
XX	PN	WO200231134-A2.	
XX	PD	18-APR-2002.	
XX	PF	12-OCT-2001; 2001WO-US031874.	
XX	PR	12-OCT-2000; 2000US-0240117P.	
XX	XX	(FERR ) FERRING BV.	
PA	PI	Qi S, Akinsanya KO, Riviere PJ, Junien J;	
XX	PI	P-PSDB; ABG61597.	
XX	DR	WPI; 2002-444178/47.	
XX	DR	P-PSDB; ABG61597.	
XX	PT	New dipeptidyl peptidase IV-related proteins and nucleic acids encoding	
XX	PT	the proteins, useful for treating e.g. fungal, bacterial, protozoan and	
XX	PT	viral infections, cancers, allergies, neurological disorders, or pain.	
XX	PS	Disclosure; Page 67-68; 113pp; English.	
XX	CC	The present invention relates to the isolation of novel human serine	
XX	CC	proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins	

CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic acids encoding them are useful for treating infections such as fungal, bacterial, protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes, precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, stroke, ulcers, asthma, allergies, cancers, migraine, vomiting, psychotic and neurological disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias. These may also be used in discovering therapeutic agents for the treatment of reproductive, inflammatory and metabolic disorders. ABK83322 -ABK83343 encode human DPRP proteins

Sequence 4309 BP; 1304 A; 790 C; 907 G; 1308 T; 0 U; 0 Other;

Query Match 97.5%; Score 1055.8; DB 6; Length 4309;  
Best Local Similarity 99.5%; Pred. No. 7.5e-313;  
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 GGAAGAAGATCCAGATCAGTGGAGTGGCTACCTTTGTTCTCCAAAGAAGATTGTATAG 60  
DB 942 GGAAGAAGATCCAGATCAGTGGAGTGGCTACCTTTGTTCTCCAAAGAAGATTGTATAG 1001  
QY 61 ATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAACAACTCCAGTGGTGGTAAATTTCT 120  
DB 1002 ATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAACAACTCCAGTGGTGGTAAATTTCT 1061  
QY 121 TAGAATTCCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCC 180  
DB 1062 TAGAATTCCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCC 1121  
QY 181 TAATGTTGAAACAAGAGGGGAGATTCATTCCTGTTATCTCTAAACAGGTACAGCAATTC 240  
DB 1122 TAATGTTGAAACAAGAGGGGAGATTCATTCCTGTTATCTCTAAACAGGTACAGCAATTC 1181  
QY 241 TAAGTCACTTTTAAGATGTCAGAAATPAATGATGATGCTGAAGGAAGATCATGTTGA 300  
DB 1182 TAAGTCACTTTTAAGATGTCAGAAATPAATGATGATGCTGAAGGAAGATCATGTTGA 1240  
QY 301 TGAAGTCAGAAGCTGGTATATTTTGAAGCACAAGACTCCCTTTAGAGCATCACCT 360  
DB 1241 TGAAGTCAGAAGCTGGTATATTTTGAAGCACAAGACTCCCTTTAGAGCATCACCT 1300  
QY 361 GTAGTGTAGTACGTTACGTTAAATCTCGAGAGGTGACAGGCTGACTGCGGTGCTACTC 420  
DB 1301 GTAGTGTAGTACGTTACGTTAAATCTCGAGAGGTGACAGGCTGACTGCGGTGCTACTC 1360  
QY 421 ACATTTCTGTCGATCAGTCACTGAGTCTTTTATAAGTAAAGTATAGTAACAGAA 480  
DB 1361 ACATTTCTGTCGATCAGTCACTGAGTCTTTTATAAGTAAAGTATAGTAACAGAA 1420  
QY 481 GAAATCCACACTGTGTGTCCTTTTACAGACTATCAAGTCTCTGAAGATGACCCAACTTGCAA 540  
DB 1421 GAAATCCACACTGTGTGTCCTTTTACAGACTATCAAGTCTCTGAAGATGACCCAACTTGCAA 1480  
QY 541 AACAAAGGAATTTTGGGCCCACTTTTGGATTTCAGAGGTCTCTTCTGCTGACTATCTCC 600  
DB 1481 AACAAAGGAATTTTGGGCCCACTTTTGGATTTCAGAGGTCTCTTCTGCTGACTATCTCC 1540  
QY 601 TCCAGAAATTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAA 660  
DB 1541 TCCAGAAATTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAA 1600  
QY 661 GCCTCATGATCTACAGCTGGAAGAAATATCTACTGCTGCTGTTTCATATATGGTG--T 718  
DB 1601 GCCTCATGATCTACAGCTGGAAGAAATATCTACTGCTGCTGTTTCATATATGGTGCTCT 1660  
QY 719 CTTAGGTCAGTGTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCGGCTTGAATACC 778  
DB 1661 CTTAGGTCAGTGTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCGGCTTGAATACC 1720  
QY 779 CTAGCCTCTCTAGGTTATGTTGTTAGTGTATGATAGCAACAGGGGATCTCTGTCACCGAGG 838

DB 1721 CTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAAACAGGGGATCTCTGTCACCGAGG 1780  
QY 839 CTTAAATTTGAAGCGCCTTTTAAATATATAAATGGGTCAAATAGAAATTCACGATCAGGTG 898  
DB 1781 CTTAAATTTGAAGCGCCTTTTAAATATATAAATGGGTCAAATAGAAATTCACGATCAGGTG 1840  
QY 899 GAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATC 958  
DB 1841 GAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATC 1900  
QY 959 CACGCTGGTCTCTATGAGAGATACCTCTCCCTGATGGCAATTAATGAGAGGTGAGATATC 1018  
DB 1901 CACGCTGGTCTCTATGAGAGATACCTCTCCCTGATGGCAATTAATGAGAGGTGAGATATC 1960  
QY 1019 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACCTCTGTGATCTTCTATATACAGATATC 1078  
DB 1961 TTCAGGTTGCTATTTGCTGGGGCCCGAGTCACCTCTGTGATCTTCTATATACAGATATC 2020  
QY 1079 ACGGA 1083  
DB 2021 ACGGA 2025  
RESULT 3  
ID ABV76411 standard; cDNA; 1821 BP.  
XX AC ABV76411;  
XX AC ABV76411;  
DT 07-FEB-2003 (first entry)  
XX Dipeptidyl peptidase IV protein 27.5-encoding cDNA.  
DE Dipeptidyl peptidase IV protein 27.5; recombinant production;  
XX Dipeptidyl peptidase IV protein 27.5; recombinant production;  
KW Stenotrophomonas maltophilia dipeptidyl peptidase IV homologue;  
KW Gene therapy; diabetes; tumour; cancer; cytostatic; gene; ss.  
XX Unidentified.  
XX Key Location/Qualifiers  
FT CDS 826..1578  
FT /\*tag= a  
FT /product= "Dipeptidyl peptidase IV protein 27.5"  
XX CN1342768-A.  
XX 03-APR-2002.  
XX 12-SEP-2000; 2000CN-00125127.  
XX 12-SEP-2000; 2000CN-00125127.  
XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.  
XX Mao Y, Xie Y;  
XX WPI; 2002-529809/57.  
XX P-PSDB; ABB99949.  
XX New dipeptidyl peptidase IV protein 27.5 and encoding polynucleotide,  
XX useful for treating diabetes and tumors.  
XX Claim 6; Page 24-25 (Disclosure); 32pp; Chinese.  
XX The invention relates to dipeptidyl peptidase IV protein 27.5 (ABB99949)  
XX and nucleic acids encoding it (ABV76411). The protein has a molecular  
XX weight of 27.5 kD and has 40% identity and 58% homology over a 227 amino  
XX acid stretch with Stenotrophomonas maltophilia dipeptidyl peptidase IV  
XX (GenBank accession number D83263). The invention also relates to a method  
XX for the recombinant production of the protein, an antagonist of the  
XX protein, and the use of the protein, gene and antagonist in therapeutic  
XX applications. Dipeptidyl peptidase IV protein 27.5 can be used in the  
XX treatment of a variety of diseases such as diabetes and tumours. The



Db 1101 TCCTCTCCGCTGACTATCTCTCTCAGAAATTTCTCTTTTGAAGTACTCTGGATTAC 1042  
Qy 640 ATTGTATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 699  
Db 1041 ATTGTATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 982  
Qy 700 GCTGTTTATATATGGTGGTCTCTCAGTGCAGTTGGTGAATATCGTTTAAAGAGTCAA 759  
Db 981 GCTGTTTATATATGGTGGTCTCTCAGTGCAGTTGGTGAATATCGTTTAAAGAGTCAA 922  
Qy 760 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGGTTGAGTAGTAGCAACAG 819  
Db 921 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGGTTGAGTAGTAGCAACAG 862  
Qy 820 GGGATCTGTGACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 879  
Db 861 GGGATCTGTGACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 802  
Qy 880 AGAAATTCAGATCAGTGAAGGACTCCCAATATCTAGCTTCTCATATGATTTTCATTGA 939  
Db 801 AGAAATTCAGATCAGTGAAGGACTCCCAATATCTAGCTTCTCATATGATTTTCATTGA 742  
Qy 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATT 999  
Db 741 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATT 682  
Qy 1000 AATGACAGAGTTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 1059  
Db 681 AATGACAGAGTTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 622  
Qy 1060 CTTCTATGATACAGGATACACGGA 1083  
Db 621 CTTCTATGATACAGGATACACGGA 598

RESULT 5

AAH99935  
ID AAH99935 standard; cDNA; 2643 BP.  
XX  
AC AAH99935;  
XX  
DT 12-APR-2002 (first entry)  
XX  
DE Coding sequence of 21953 human prollyl oligopeptidase.  
XX  
KW 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;  
KW cardiovascular disease; autoimmune disease; atopic allergy;  
KW neuronal disorder; vascular disorder; prostate disorder; Cystostatic;  
KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;  
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;  
KW Grave's disease; neuronal disorder; demyelinating disease; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200179473-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 11-APR-2001; 2001WO-US040483.  
XX  
PR 18-APR-2000; 2000US-0197508P.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Meyers RA, Williamson M;  
XX  
XX WPI; 2002-034353/04.  
DR P-PSDB; AAG78415.  
XX  
PT New polypeptides 21953, member of human prollyl oligopeptidase family,  
PT useful as diagnostic targets and therapeutic agents for controlling  
PT cancer, lymphoma and leukemia.  
XX

PS Claim 2; Page; 121pp; English.  
XX  
This invention relates to an isolated 21953 human prollyl oligopeptidase.  
CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,  
CC antithyroid, dermatological, antipsoriatic, antiasthmatic,  
CC ophthalmological, antiinflammatory, nootropic, antiparkinsonian,  
CC anticonvulsant, gynecological, vasorelaxant, antianginal, cardiac,  
CC antithrombotic, anorectic and metabolic in its action. Uses include  
CC gene therapy, expression or activity of 21953 protein modulator, it is  
CC useful for identifying a compound which binds to it and can be used in  
CC preventing, treating or detecting a cellular proliferative or  
CC differentiative disorder. The 21953 molecules can act as novel diagnostic  
CC targets and therapeutic agents for controlling disorders associated with  
CC the aberrant activity or degradation of peptide hormones e.g., disorders  
CC associated with cell differentiation and proliferation such as cancer,  
CC immune function, reproductive, neurological and cardiovascular function.  
CC The 21953 molecules are thus useful for treating and preventing cellular  
CC proliferative and differentiative disorders, haematopoietic neoplastic  
CC disorders, immune disorders such as autoimmune diseases, diabetes  
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,  
CC neuronal disorders, demyelinating diseases, vascular disorders and  
CC metabolism or pain disorders. This sequence represents the cDNA encoding  
CC sequence of 21953 human prollyl oligopeptidase. This sequence represents  
CC the coding sequence of 21953, being the sequence in between the start and  
CC the stop codon of the sequence represented in AAH99934  
XX  
SQ Sequence 2643 BP; 800 A; 514 C; 585 G; 744 T; 0 U; 0 Other;  
Query Match 73.1%; Score 791.2; DB 6; Length 2643;  
Best Local Similarity 99.0%; Pred. No. 1.1e-231;  
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Qy 280 TGAAGGAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339  
Db 1524 TGGATCTAATATCAAGTTCAGTGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1583  
Qy 340 CTCCTCTTAGAGCATCACCTGTACGTAGTCAGTACGTAAATCTGGAGAGGTGACAG 399  
Db 1584 CTCCTCTTAGAGCATCACCTGTACGTAGTCAGTACGTAAATCTGGAGAGGTGACAG 1643  
Qy 400 GCTGACTGACCGTGGCTACTCACATCTTGCTGCATCAGTCAGCACTGTCTTTAT 459  
Db 1644 GCTGACTGACCGTGGCTACTCACATCTTGCTGCATCAGTCAGCACTGTCTTTAT 1703  
Qy 460 AAGTAAGTATAGTAACCCAGAGAAATCCACTGTGTCCCTTTACAAGCTATCAAGTCC 519  
Db 1704 AAGTAAGTATAGTAACCCAGAGAAATCCACTGTGTCCCTTTACAAGCTATCAAGTCC 1763  
Qy 520 TGAAGATGACCCCACTTGCAAAACAAAGGAATTTTGGGCCACCATTTGGATTCAGCAGG 579  
Db 1764 TGAAGATGACCCCACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGG 1823  
Qy 580 TCCTCTTCCTGACTATATCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGGATTAC 639  
Db 1824 TCCTCTTCCTGACTATATCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGGATTAC 1883  
Qy 640 ATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 699  
Db 1884 ATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAATAATCTCTACTGT 1943  
Qy 700 GCTGTTTATATATGGTGGTCTCTCAGTGCAGTTGGTGAATATCGTTTAAAGAGTCAA 759  
Db 1944 GCTGTTTATATATGGTGGTCTCTCAGTGCAGTTGGTGAATATCGTTTAAAGAGTCAA 2003  
Qy 760 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGGTTGAGTAGTAGCAACAG 819  
Db 2004 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTTGGTTGAGTAGTAGCAACAG 2063  
Qy 820 GGGATCTGTGACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 879  
Db 2064 GGGATCTGTGACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 2123  
Qy 880 AGAAATTCAGATCAGTGAAGGACTCCCAATATCTAGCTTCTCATATGATTTTCATTGA 939

DB 2124 AGAATTCAGCATCAGGTGGAGGATCCATATCTAGCTTCTCGATGATTTTCATGA 2183  
QY 940 CTTAGATCGGTGGGATCCACGGCTGCTTATGAGGATACCTCTCCCTGATGGCAAT 999  
DB 2184 CTTAGATCGGTGGGATCCACGGCTGCTTATGAGGATACCTCTCCCTGATGGCAAT 2243  
QY 1000 AATCAGAGTCTAGATCTTTCAGGTTCTTATGCTGGGGCCCCAGTCACTCTGTGGAT 1059  
DB 2244 AATCAGAGTCTAGATCTTTCAGGTTCTTATGCTGGGGCCCCAGTCACTCTGTGGAT 2303  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
DB 2304 CTTCTATGATACAGGATACACGGA 2327

RESULT 6  
ID ABX12255  
XX ABX12255 standard; cDNA; 2649 BP.  
AC  
XX  
DT 19-MAY-2003 (first entry)  
XX  
DE cDNA encoding human serine protease HIPHUM46.

XX Human; ss; gene; HIPHUM46; serine protease; gene therapy; osteoarthritis;  
KW serine protease activity modulation; dipeptidyl peptidase activity; prophy-

XX musculoeskeletal disease; Hepatitis B virus infection; myotonic dystrophy;  
KW Alzheimer's disease; paraspranuclear palsy; Huntington's disease;  
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;  
KW irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;  
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;  
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;  
XX multiple sclerosis; chromosome 15q21-q22.

OS Homo sapiens.

EH Key Location/Qualifiers  
FT CDS 1..2649  
FT /\*tag= a  
FT /product= "HIPHUM46"  
FT /note= "Serine protease"

XX GB2374869-A.  
XX 30-OCT-2002.  
XX 22-JAN-2002; 2002GB-00001404.  
XX 23-JAN-2001; 2001GB-00001760.

XX (GLAX ) GLAXO GROUP LTD.

XX Edbrooke MR, Lewis AP;

XX WPI; 2003-150703/15.

XX P-PSDB; ABU07720.

PT Identifying modulators of serine protease activity useful for treating  
PT musculoeskeletal diseases, by contacting cell expressing a novel serine  
PT protease polypeptide with a compound and monitoring serine protease  
PT activity.

PS Claim 12; Page 22-26; 38pp; English.

XX The invention relates to a method of identifying a substance that  
CC modulates serine protease activity, comprising contacting a cell such as  
CC a neuronal cell, lung cell, intestinal cell or a cell infected with a  
CC virus, expressing a serine protease polypeptide (HIPHUM 46), or its  
CC variant having dipeptidyl peptidase activity, or a serine protease  
CC isolated from the cell with a test substance and monitoring for serine  
CC protease activity. The method is useful for identifying a substance that

CC modulates serine protease activity. A modulator of the serine protease is  
CC useful in the manufacture of a medicament for treatment or prophylaxis of  
CC a musculoeskeletal disease e.g. osteoarthritis, Hepatitis B virus  
CC infection, Alzheimer's disease, paraspranuclear palsy, myotonic  
CC dystrophy, Huntington's disease or amyotrophic lateral sclerosis.  
CC Additional disease that may be treated using modulators of the serine  
CC protease include malabsorption syndromes, irritable bowel syndrome, lung  
CC disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,  
CC rectal polyps, small bowel tumours, colorectal tumours, anaemia,  
CC dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple  
CC sclerosis. The present sequence represents cDNA encoding the human serine  
CC protease HIPHUM46 which is located on chromosome 15q21-q22

XX Sequence 2649 BP; 803 A; 514 C; 585 G; 747 T; 0 U; 0 Other;

Query Match 73.1%; Score 791.2; DB 8; Length 2649;  
Best Local Similarity 99.0%; Pred. No. 1.1e-231;  
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAAGGATCATAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGA 339  
DB 1527 TGGATCTAATATCCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCACCAAGA 1586  
QY 340 CTCCTCTTATAGAGCATCCTGTACGTACGTACGTACGTACGTACGTACGTACGTACGTAC 399  
DB 1597 CTCCTCTTATAGAGCATCCTGTACGTACGTACGTACGTACGTACGTACGTACGTAC 1646  
QY 400 GCTGACTGACCGGTGGTACTACATCTTCTGCTGCTACGTACGTACGTACGTACGTACGT 459  
DB 1647 GCTGACTGACCGGTGGTACTACATCTTCTGCTGCTACGTACGTACGTACGTACGTAC 1706  
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 519  
DB 1707 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 1766  
QY 520 TGAAGATGACCCCAACTTGCAAAACAAAGAAATTTTGGGCCCACTTTTGGATTGACGAG 579  
DB 1767 TGAAGATGACCCCAACTTGCAAAACAAAGAAATTTTGGGCCCACTTTTGGATTGACGAG 1826  
QY 580 TCCTCTTCTGACTATACCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTAC 639  
DB 1827 TCCTCTTCTGACTATACCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTAC 1886  
QY 640 ATTGTATGGGATGCTCTCAAGGCTCATGATCTACAGCTCGGAAAGAAATATCTTACTGT 699  
DB 1887 ATTGTATGGGATGCTCTCAAGGCTCATGATCTACAGCTCGGAAAGAAATATCTTACTGT 1946  
QY 700 GCTGTTTCATATATGTTGGTCTCCTCAGGTGAGTTGGTGAATTAATCGGTTTAAAGGAGTCAA 759  
DB 1947 GCTGTTTCATATATGTTGGTCTCCTCAGGTGAGTTGGTGAATTAATCGGTTTAAAGGAGTCAA 2006  
QY 760 GTATTTCCGCTTGAATACCTTAGCCCTCTCTAGGTTTATGTTGGTTAGTAGACACAG 819  
DB 2007 GTATTTCCGCTTGAATACCTTAGCCCTCTCTAGGTTTATGTTGGTTAGTAGACACAG 2066  
QY 820 GGGATCCTGTACCGAGGGCTTAAATTTGAAGCGGCTTAAATATAAATAAATGGGTCAAT 879  
DB 2067 GGGATCCTGTACCGAGGGCTTAAATTTGAAGCGGCTTAAATATAAATAAATGGGTCAAT 2126  
QY 880 AGAATTTGACATCAGTCTGAGAGGACTCCAAATATCTAGCTTCTCGATGATTTTCATGA 939  
DB 2127 AGAATTTGACATCAGTCTGAGAGGACTCCAAATATCTAGCTTCTCGATGATTTTCATGA 2186  
QY 940 CTTAGATCGTGTGGGATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAAT 999  
DB 2187 CTTAGATCGTGTGGGATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAAT 2246  
QY 1000 AATGCAAGGTCTAGATATCTTTCAGGTTGCTTATTCCTGGGGCCCCAGTCACTCTGTGGAT 1059  
DB 2247 AATGCAAGGTCTAGATATCTTTCAGGTTGCTTATTCCTGGGGCCCCAGTCACTCTGTGGAT 2306  
QY 1060 CTTCTATGATACAGGATACACGGA 1083

Db 2307 CTCTATGATACAGGATACACGGA 2330

RESULT 7  
ADV43981  
ID ADV43981 standard; cDNA; 2649 BP.  
XX  
AC  
XX  
XX  
DT 10-MAR-2005 (first entry)  
XX  
DE Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 1609.  
XX  
KW microarray; psychoneuroendocrinimmune; chronic fatigue;  
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;  
KW cancer; neoplasm; infection; expressed sequence tag; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2004108899-A2.  
XX  
PD 16-DEC-2004.  
XX  
PF 04-JUN-2004; 2004WO-US017686.  
XX  
PR 04-JUN-2003; 2003US-0475915P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Nicholson A, Vernon SD;  
XX  
XX WPI; 2005-031682/03.  
DR  
XX  
XX New microarray comprising probes for genes involved in  
PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a  
PT condition associated with PNI activity, e.g., inflammatory or infectious  
PT diseases.  
XX  
PS Claim 1; SEQ ID NO 1609; 254pp; English.  
XX  
XX The invention relates to a new microarray which comprises probes for  
CC genes involved in psychoneuroendocrinimmune (PNI) activity. The  
CC microarray is useful in diagnosing a condition associated with PNI  
CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,  
CC cancer and infection. The present sequence represents a  
CC psychoneuroendocrinimmune gene expressed sequence tag. Note the  
CC specificatio mentions SEQ ID NO of up to 3314 but only sequences up to  
CC SEQ ID NO 1829 are provided.  
XX  
SQ Sequence 2649 BP; 804 A; 514 C; 585 G; 746 T; 0 U; 0 Other;

Query Match 73.1%; Score 791.2; DB 14; Length 2649;  
Best Local Similarity 99.0%; Pred. No. 1.1e-231;  
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAGGAAGGATCATAGTTCAGTCAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339  
DB 1527 TGGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1586

QY 340 CTCCTCTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAG 399  
DB 1587 CTCCTCTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAG 1646

QY 400 GCTGACTGACCGTGGCTACTCACATTCTGCTGCATCAGTCAGTGTGACTCTTTTAT 459  
DB 1647 GCTGACTGACCGTGGCTACTCACATTCTGCTGCATCAGTCAGTGTGACTCTTTTAT 1706

QY 460 AAGTAAGTATAGTACCAAGGAATCCACACTGTGTGCTCTTACAAAGCTATCAAGTCC 519  
DB 1707 AAGTAAGTATAGTACCAAGGAATCCACACTGTGTGCTCTTACAAAGCTATCAAGTCC 1766

QY 520 TGAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACGAGG 579

Db 1767 TGAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACGAGG 1826

QY 580 TCCTCTTCTGACTATACTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCTGATTTAC 639

Db 1827 TCCTCTTCTGACTATACTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCTGATTTAC 1886

QY 640 ATTGTATGGGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 699

Db 1887 ATTGTATGGGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 1946

QY 700 GCTGTTTCATATATGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 759

Db 1947 GCTGTTTCATATATGATGCTCTTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 2006

QY 760 GTATTTCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTAGTATGACACACAG 819

Db 2007 GTATTTCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGTTAGTAGTATGACACACAG 2066

QY 820 GGGATCCTGTACACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAAT 879

Db 2067 GGGATCCTGTACACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAAT 2126

QY 880 AGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGA 939

Db 2127 AGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGA 2186

QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 999

Db 2187 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 2246

QY 1000 AATGAGAGGTCAGATATCTTCAAGGTTGCTATGCTGGGGCCCACTCACTCTGTGGAT 1059

Db 2247 AATGAGAGGTCAGATATCTTCAAGGTTGCTATGCTGGGGCCCACTCACTCTGTGGAT 2306

QY 1060 CTTCTATGATACAGGATACACGGA 1083

Db 2307 CTTCTATGATACAGGATACACGGA 2330

RESULT 8  
ADL13374  
ID ADL13374 standard; cDNA; 2797 BP.  
XX  
AC ADL13374;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
XX Human steroid-induced C3A liver cell cDNA #1103.  
XX  
KW ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;  
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.  
XX  
OS Homo sapiens.  
XX  
PN US6673549-B1.  
XX  
PD 06-JAN-2004.  
XX  
PF 12-OCT-2001; 2001US-00976594.  
XX  
PR 12-OCT-2000; 2000US-0240409P.  
XX  
PA (INCY-) INCYTE CORP.  
XX  
PI Furness LM, Buchbinder JL;  
XX  
XX WPI; 2004-068610/07.  
XX  
XX Combination useful for preparing a composition for treating liver  
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,  
PT comprises cDNAs that are differentially expressed in response to steroid  
PT treatment.  
XX



PS	Claim 1; SEQ ID NO 1103; 141pp; English.	
XX	The invention relates to a combination comprising cDNAs that are	
CC	differentially expressed in response to steroid treatment. Also included	
CC	are the following: a high throughput method for using a cDNA to detect	
CC	differential expression of nucleic acids in a sample; and a high	
CC	throughput method of screening molecules or compounds to identify a	
CC	ligand that specifically binds a cDNA. The sample is from a subject with	
CC	Wilson disease and comparison of a standard defines a stage of that	
CC	disease. The high throughput method of screening molecules or compounds	
CC	to identify a ligand that specifically binds a cDNA comprises: combining	
CC	the combination with molecules or compounds under conditions to allow	
CC	specific binding; and detecting specific binding between each cDNA and at	
CC	least one molecule or compound. The molecules or compounds are regulatory	
CC	proteins. The combination is useful for preparing a composition for	
CC	treating liver disorders associated with steroid therapy, e.g., cirrhosis	
CC	or hepatitis. The present sequence represents a human cDNA which is	
CC	differentially expressed in steroid-induced C3A liver cells. Note: The	
CC	sequence data for this patent did not form part of the printed	
CC	specification but was obtained in electronic format directly from USPTO	
CC	at <a href="http://seqdata.uspto.gov/sequence.html">seqdata.uspto.gov/sequence.html</a> .	
XX		
SQ	Sequence 2797 BP; 866 A; 551 C; 595 G; 785 T; 0 U; 0 Other;	
	Query Match 73.1%; Score 791.2; DB 12; Length 2797;	
	Best Local Similarity 99.0%; Pred. No. 1.1e-231;	
	Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
QY	280 TGAAGGAAGGATCATAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGA 339	
DB	1411 TGGATCTTAATCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGA 1470	
QY	340 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAAG 399	
DB	1471 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACAAAG 1530	
QY	400 GCTGACTGACCGTGGCTACTACATCTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTAT 459	
DB	1531 GCTGACTGACCGTGGCTACTACATCTTCTGCTGCATCAGTCAGCACTGTGACTTCTTTAT 1590	
QY	460 AAGTAAGTATAGTAACAGGAAGATCACACTGTGTCCCTTTTCAAGCTATCAAGTCC 519	
DB	1591 AAGTAAGTATAGTAACAGGAAGATCACACTGTGTCCCTTTTCAAGCTATCAAGTCC 1650	
QY	520 TGAAGATGACCACTTGCRAAACCAAGGAATTTTGGGCCACCATTTTGGATTCCAGCAGG 579	
DB	1651 TGAAGATGACCACTTGCRAAACCAAGGAATTTTGGGCCACCATTTTGGATTCCAGCAGG 1710	
QY	580 TCCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTAC 639	
DB	1711 TCCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTAC 1770	
QY	640 ATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAANAATATCCTACTGT 699	
DB	1771 ATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAANAATATCCTACTGT 1830	
QY	700 GCTGTTTCATATATGGTGGTCTCTCAGTGCAGTTGGTGAATAATTCGGTTTAAAGAGTCAA 759	
DB	1831 GCTGTTTCATATATGGTGGTCTCTCAGTGCAGTTGGTGAATAATTCGGTTTAAAGAGTCAA 1890	
QY	760 GTATTTCGGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTAGTGATAGACAACAG 819	
DB	1891 GTATTTCGGCTTGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTAGTGATAGACAACAG 1950	
QY	820 GGGATCTGTCCAGCGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAAAT 879	
DB	1951 GGGATCTGTCCAGCGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAAAT 2010	
QY	880 AGAAATTCACCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 939	
DB	2011 AGAAATTCACCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATGA 2070	
QY	940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATT 999	

---

DB	2071 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATT 2130	
QY	1000 AATGCAGAGGTTCAGATATCTTCAGGGTTCCTATTGTGGGGCCCCAGTCACTCTGTGGAT 1059	
DB	2131 AATGCAGAGGTTCAGATATCTTCAGGGTTCCTATTGTGGGGCCCCAGTCACTCTGTGGAT 2190	
QY	1060 CTTCTATGATACAGGATACACGGA 1083	
DB	2191 CTTCTATGATACAGGATACACGGA 2214	
RESULT 9		
ID	ACA92425 standard; DNA; 2929 BP.	
XX	ACA92425;	
XX	DT 15-JUL-2003 (first entry)	
XX	XX DNA encoding human PMMM-10.	
DE		
XX	Human; protein modification and maintenance molecule; PMMM; cancer;	
XX	cell proliferation disorder; atherosclerosis, neurological disorder;	
XX	epilepsy; Huntington's disease; stroke; immune disorder; allergy;	
XX	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;	
XX	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;	
XX	infection; cytostatic; antiarteriosclerotic; anticonvulsant; neurotic;	
XX	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary;	
XX	antiinflammatory; thymimetic; gene; ds.	
OS	Homo sapiens.	
XX	WO2003031939-A2.	
XX	17-APR-2003.	
XX	11-OCT-2002; 2002WO-US032850.	
XX	12-OCT-2001; 2001US-0329689P.	
XX	25-OCT-2001; 2001US-0335703P.	
PR	09-NOV-2001; 2001US-034887P.	
PR	28-NOV-2001; 2001US-0334145P.	
PR	06-DEC-2001; 2001US-0337451P.	
PR	14-DEC-2001; 2001US-0340584P.	
XX	(INCY-) INCYTE GENOMICS INC.	
XX	Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;	
PI	Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;	
PI	Sprague WW, Hafalia AJA, Chawla NK, Lehr-Nason PM, Kable AE, Yue H;	
PI	Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;	
PI	Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W,	
XX	WPI; 2003-430274/40.	
DR	P-PSDB; ABU92030.	
XX	New human protein modification and maintenance molecules (PMMM), useful	
PT	for diagnosing, treating and preventing diseases or conditions associated	
PT	with the aberrant PMMM expression e.g. cancer, atherosclerosis, or	
PT	infections.	
XX	Claim 5; Page 289; 311pp; English.	
XX	The present invention relates to the isolation of human protein	
CC	modification and maintenance molecules (PMMM), and the polynucleotide	
CC	sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM	
CC	-1 to PMMM-40) are disclosed. The sequences of the invention are useful	
CC	for diagnosing a condition or disease associated with the expression of	
CC	PMMM in a subject, preparing a polyclonal or monoclonal antibody, and	
CC	generating an expression profile of a sample containing the	
CC	polynucleotides. The diseases or conditions associated with decreased	
CC	expression or overexpression of PMMM are cell proliferation disorders	



CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
CC Huntington's disease, stroke), immune/inflammatory disorders (e.g. AIDS,  
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
CC syndrome), gastrointestinal or epithelial disorders, and infections. The  
CC PMM polypeptides or their fragments are useful in screening compounds  
CC for effectiveness as agonists or antagonists of the polypeptides, or in  
CC altering the expression of the target polynucleotide and compounds that  
CC specifically bind to, or modulate the activity of the polypeptide.  
CC ACA92416-ACA92455 encode the human PMM polypeptides of the invention  
XX  
SQ Sequence 2929 BP; 864 A; 606 C; 661 G; 798 T; 0 U; 0 Other;  
Query Match 73.1%; Score 791.2; DB 10; Length 2929;  
Best Local Similarity 99.0%; Pred. No. 1.1e-231;  
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 280 TGAAGGAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339  
Db |||||  
1555 TGGATCTAATATCCAAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1614  
QY 340 CTCCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTCGAGAGGTCACAAG 399  
Db |||||  
1615 CTCCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTCGAGAGGTCACAAG 1674  
QY 400 GCTGACTGACCGTGGCTACTCATCATTTCTTGCTGCATCAGTCAGCACTGTGACTCTTTTAT 459  
Db |||||  
1675 GCTGACTGACCGTGGCTACTCATCATTTCTTGCTGCATCAGTCAGCACTGTGACTCTTTTAT 1734  
QY 460 AAGTAAGTATAGTACCAAGAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCC 519  
Db |||||  
1735 AAGTAAGTATAGTACCAAGAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCC 1794  
QY 520 TGAAGATGACCCCACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 579  
Db |||||  
1795 TGAAGATGACCCCACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 1854  
QY 580 TCCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGAATTTAC 639  
Db |||||  
1855 TCCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGAATTTAC 1914  
QY 640 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGAGAAAGAAATATCCTACTGT 699  
Db |||||  
1915 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGAGAAAGAAATATCCTACTGT 1974  
QY 700 GCTGTTTCATATATGGTGGTCTCTCAGGTGCAATTCGGTTTAAAGAGTCAA 759  
Db |||||  
1975 GCTGTTTCATATATGGTGGTCTCTCAGGTGCAATTCGGTTTAAAGAGTCAA 2034  
QY 760 GTATTTCCGTTGAATACCTAGCCTCTAGGTTATGTTGTTAGTATAGACACAG 819  
Db |||||  
2035 GTATTTCCGTTGAATACCTAGCCTCTAGGTTATGTTGTTAGTATAGACACAG 2094  
QY 820 GGAATCTGTCCAGCGGGCTTAAATTTGAAGCGCGCTTAAATATAAAATGGGTCAAAT 879  
Db |||||  
2095 GGAATCTGTCCAGCGGGCTTAAATTTGAAGCGCGCTTAAATATAAAATGGGTCAAAT 2154  
QY 880 AGAATATGACCATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGA 939  
Db |||||  
2155 AGAATATGACCATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTCAATGA 2214  
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATT 999  
Db |||||  
2215 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATT 2274  
QY 1000 AATGACAGGTCAGATATCTTCAGGGTTCATATGCTGGGGCCCCAGTCATCTGTGGAT 1059  
Db |||||  
2275 AATGACAGGTCAGATATCTTCAGGGTTCATATGCTGGGGCCCCAGTCATCTGTGGAT 2334  
QY 1060 CTTCTATGATACAGGATACAGGA 1083  
Db |||||  
2335 CTTCTATGATACAGGATACAGGA 2358

RESULT 10  
ID ABK12892 standard; cDNA; 3106 BP.  
XX  
AC ABK12892;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Human protease PRTS-9 cDNA sequence.  
XX  
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
KW cell proliferative disorder; developmental disorder; epilepsy;  
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
KW reproductive disorder; endometriosis; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 203..2851  
FT /\*tag= a  
ET /product= "Human protease PRTS-9"  
XX  
PN WO200198468-A2.  
XX  
PD 27-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019178.  
XX  
PR 16-JUN-2000; 2000US-0212336P.  
PR 22-JUN-2000; 2000US-0213955P.  
PR 29-JUN-2000; 2000US-0215396P.  
PR 07-JUL-2000; 2000US-0216821P.  
PR 14-JUL-2000; 2000US-0218946P.  
XX  
(INCY-) INCYTE GENOMICS INC.  
XX  
PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;  
PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
PI Walla NK, Yao MG, Lu DM, Patterson C, Tang YT, Walsh RT;  
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
PI Kallick DA;  
XX  
DR WPI; 2002-090437/12.  
DR P-PSDB; AAU74749.  
XX  
PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in  
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative  
PT (e.g. cancer) disorders.  
XX  
PS Claim 5; Page 166-167; 177pp; English.  
XX  
CC The present invention relates to twenty one new human proteases, referred  
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the  
CC invention are useful in the diagnosis, treatment and prevention of  
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's  
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial  
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency  
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.  
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,  
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's  
CC disease and reproductive e.g. infertility and endometriosis disorders.  
CC Numerous other examples of each disorder are given in the specification.  
CC The present nucleic acid sequence encodes the human protease PRTS-9  
CC protein of the invention  
XX  
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;  
Query Match 73.1%; Score 791.2; DB 6; Length 3106;  
Best Local Similarity 99.0%; Pred. No. 1.2e-231;  
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAGGATCATAGTTCATAGTCAGAGGCTGGTATATATTTTGAAGGCACCAAGA 339  
Db 1729 TGGATCTAATATCCAGTTCATAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1788  
QY 340 CTCGCCCTTTAGAGATCACCTGTACGTAGTTCAGTTCAGTAAATCTCGAGAGGTGACAAG 399  
Db 1789 CTCGCCCTTTAGAGATCACCTGTACGTAGTTCAGTTCAGTAAATCTCGAGAGGTGACAAG 1848  
QY 400 GCTGACTGACCCGTGGCTACTCAATTCTTGTGCTGATCAGTCAGCACTGTGACTTCTTTAT 459  
Db 1849 GCTGACTGACCCGTGGCTACTCAATTCTTGTGCTGATCAGTCAGCACTGTGACTTCTTTAT 1908  
QY 460 AAGTAAGTATAGTAAACAGAGGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 519  
Db 1909 AAGTAAGTATAGTAAACAGAGGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 1968  
QY 520 TGAAGATGACCCCACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 579  
Db 1969 TGAAGATGACCCCACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 2028  
QY 580 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTCAC 639  
Db 2029 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTCAC 2088  
QY 640 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGGAAGAAATATCTCTACTGT 699  
Db 2089 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGGAAGAAATATCTCTACTGT 2148  
QY 700 GCTGTTTCATATATGGTGGTCTCTCAGTGCAGTTGGTGTATATCGTTTAAAGAGTCAA 759  
Db 2149 GCTGTTTCATATATGGTGGTCTCTCAGTGCAGTTGGTGTATATCGTTTAAAGAGTCAA 2208  
QY 760 GTATTTCGGTTCGAATACCTAGCTCTCTAGGTTATGTGTGTAGTGTAGTACACACAG 819  
Db 2209 GTATTTCGGTTCGAATACCTAGCTCTCTAGGTTATGTGTGTAGTGTAGTACACACAG 2268  
QY 820 GGGATCTGTACCCAGGCTTAAATTTGAAGCGCTTTAAATATAAATTTGGTCAAT 879  
Db 2269 GGGATCTGTACCCAGGCTTAAATTTGAAGCGCTTTAAATATAAATTTGGTCAAT 2328  
QY 880 AGAATTTGACGATCAGGTGGAAGGACTCCAAATCTAGCTTCTGATATGATTTTCATGA 939  
Db 2329 AGAATTTGACGATCAGGTGGAAGGACTCCAAATCTAGCTTCTGATATGATTTTCATGA 2388  
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGAGATACCTCTCCCTGATGGCAT 999  
Db 2389 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGAGATACCTCTCCCTGATGGCAT 2448  
QY 1000 AATCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGCCCCCAGTCACTCTGTGGAT 1059  
Db 2449 AATCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGCCCCCAGTCACTCTGTGGAT 2508  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
Db 2509 CTTCTATGATACAGGATACACGGA 2532

## RESULT 11

AAC85694  
ID AAC85694 standard; cDNA; 3120 BP.XX  
AC AAC85694;XX  
DT 29-JUN-2001 (first entry)XX  
DE Nucleotide sequence of human DPP8.XX  
KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;  
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;  
KW growth hormone deficiency; glucose level; mucosal regeneration;  
KW Non-insulin dependent diabetes mellitus; glucose intolerance;  
KW immunosuppression; ss.

XX OS Homo sapiens.  
XX Key Location/Qualifiers  
FH CDS 214..2862  
FT /\*tag= a  
FT /product= "Human DPP8"  
PN WO200119866-A1.  
XX 22-MAR-2001.  
XX 11-SEP-2000; 2000WO-AU001085.  
XX 10-SEP-1999; 99AU-00002762.  
XX 18-FEB-2000; 2000AU-00005709.  
XX (UNSY ) UNIV SYDNEY.  
XX PA Abbott CA, Gorell MD;  
XX PI P-PSDB; AAB47187.  
XX DR WPI; 2001-281520/29.  
XX DR P-PSDB; AAB47187.  
XX PT New human dipeptidyl aminopeptidase (DPP8) useful for cleaving  
XX substrates, identifying inhibitors of DPP8 catalytic activity which have  
XX therapeutic uses, and for detecting activated T cells.  
XX Claim 16; Fig 2; 78pp; English.  
XX This sequence encodes human dipeptidyl aminopeptidase (DPP8). DPP8 has  
XX substrate specificity for H-Ala-Pro-PNA, H-Gly-Pro-PNA and H-Arg-Pro-PNA.  
XX Therefore it is a prolyl oligopeptidase and a dipeptidyl peptidase.  
XX because it is capable of hydrolysing the peptide bond C-terminal to Pro  
XX in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is  
XX useful for cleaving a substrate, and for detecting an activated T cell  
XX which involves measuring the level of DPP8 gene expression in a T cell.  
XX The level of DPP8 expression is detected by detecting the amount of DPP8  
XX RNA in the cell. It is also useful for identifying a molecule capable of  
XX inhibiting the cleavage of the substrate by DPP8. Molecules identified as  
XX inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,  
XX growth hormone deficiency, lowering glucose levels in non-insulin  
XX dependent diabetes mellitus and other disorders involving glucose  
XX intolerance, enhancing mucosal regeneration and as immunosuppressants  
XX SQ Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;  
Query Match 73.1%; Score 791.2; DB 4; Length 3120;  
Best Local Similarity 59.0%; Pred. No. 1.2e-231;  
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 280 TGAAGGAGGATCATAGTTCATAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339  
Db 1740 TGGATCTAATATCCAGTTCATAGTTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799  
QY 340 CTCGCCCTTTAGAGATCACCTGTACGTAGTTCAGTTCAGTAAATCTCGAGAGGTGACAAG 399  
Db 1800 CTCGCCCTTTAGAGATCACCTGTACGTAGTTCAGTTCAGTAAATCTCGAGAGGTGACAAG 1859  
QY 400 GCTGACTGACCCGTGGCTACTCAATTCTTGTGCTGATCAGTCAGCACTGTGACTTCTTTAT 459  
Db 1860 GCTGACTGACCCGTGGCTACTCAATTCTTGTGCTGATCAGTCAGCACTGTGACTTCTTTAT 1919  
QY 460 AAGTAAGTATAGTAAACAGAGGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 519  
Db 1920 AAGTAAGTATAGTAAACAGAGGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 1979  
QY 520 TGAAGATGACCCCACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 579  
Db 1980 TGAAGATGACCCCACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGG 2039  
QY 580 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTCAC 639







PS Disclosure; Page 65-66; 113pp; English.

XX The present invention relates to the isolation of novel human serine  
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins  
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic  
CC acids encoding them are useful for treating infections such as fungal,  
CC bacterial, protozoan and viral infections, particularly infections caused  
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,  
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's  
CC disease, acute heart failure, hypotension, hypertension, urinary  
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,  
CC allergies, cancers, migraine, vomiting, psychotic and neurological  
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.  
CC These may also be used in discovering therapeutic agents for the  
CC treatment of reproductive inflammatory and metabolic disorders. ABK83322  
CC -ABK83343 encode human DPRP proteins

XX Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;

Query Match 71.9%; Score 779.2; DB 6; Length 4829;  
Best Local Similarity 98.8%; Pred. No. 7.2e-228;  
Matches 796; Conservative 0; Mismatches 8; Indels 2; Gaps 1;  
QY 280 TGAAGGAGGATCATAGTTGATGAGTCAGAGGCTGGGTATATTTGAAGGCCACCAAGA 339  
DB 1740 TGGATCTAAATCAAGTTGATGAGTCAGAGGCTGGGTATATTTGAAGGCCACCAAGA 1799  
QY 340 CTCCTCTTTAGACATCACTGTPACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAG 399  
DB 1800 CTCCTCTTTAGACATCACTGTPACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAG 1859  
QY 400 GCTGACTGACCGTGGGTACTCACAATCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTTAT 459  
DB 1860 GCTGACTGACCGTGGGTACTCACAATCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTTAT 1919  
QY 460 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTCCCTTTACAGCTATCAAGTCC 519  
DB 1920 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTCCCTTTACAGCTATCAAGTCC 1979  
QY 520 TGAAGATGACCCAACTTGCAGAAAGGAAATTTGGGCCACCAATTTGGATTTCAGCAGG 579  
DB 1980 TGAAGATGACCCAACTTGCAGAAAGGAAATTTGGGCCACCAATTTGGATTTCAGCAGG 2039  
QY 580 TCCTCTTCTGACTACTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTAC 639  
DB 2040 TCCTCTTCTGACTACTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGATTTAC 2099  
QY 640 ATTGTATGGATGCTCTACAGCCTCATGATCTACAGCTGGGAAAGAAATATCCTACTGT 699  
DB 2100 ATTGTATGGATGCTCTACAGCCTCATGATCTACAGCTGGGAAAGAAATATCCTACTGT 2159  
QY 700 GCTGTTTCATATAGTGG--TCCTCAGGTGAGTGGTGAATATATCGGTTTAAAGGAGTC 757  
DB 2160 GCTGTTTCATATAGTGGTCTCCTCAGGTGAGTGGTGAATATATCGGTTTAAAGGAGTC 2219  
QY 758 AGTATTTCCGTTGAATACCTAGCCTCTCTAGGTTATGTGTTTGTAGTGATAGACAAC 817  
DB 2220 AGTATTTCCGTTGAATACCTAGCCTCTCTAGGTTATGTGTTTGTAGTGATAGACAAC 2279  
QY 818 AGGGATCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAA 877  
DB 2280 AGGGATCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAA 2339  
QY 878 ATAGAAATTCAGATCAGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATT 937  
DB 2340 ATAGAAATTCAGATCAGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCATT 2399  
QY 938 GACTTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCA 997  
DB 2400 GACTTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCA 2459  
QY 998 TTAATGCAGATCATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGG 1057

Db 2460 TTAATGCAGAGGTCAAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGG 2519  
QY 1058 ATCTTCTATGATACAGGATACACCGA 1083  
DB 2520 ATCTTCTATGATACAGGATACACCGA 2545

Search completed: May 4, 2006, 09:09:39  
Job time : 611.925 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 02:59:51 ; Search time 4056.93 Seconds  
(without alignments)  
12489.839 Million cell updates/sec

Title: US-10-825-632-8  
Perfect score: 1083  
Sequence: 1 ggaagaagatgccagatcag.....tatgatacaggatacacgga 1083

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: gb\_est1.\*
- 2: gb\_est2.\*
- 3: gb\_est3.\*
- 4: gb\_est4.\*
- 5: gb\_est5.\*
- 6: gb\_est6.\*
- 7: gb\_est7.\*
- 8: gb\_est8.\*
- 9: gb\_est9.\*
- 10: gb\_est10.\*
- 11: gb\_est11.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	794.6	73.4	2649	10 AY411615	AY411615 Homo sapi
2	789.8	72.9	2649	10 AY411616	AY411616 Pan trogl
3	789.2	72.9	1042	3 BM557438	BM557438 AGENCOURT
4	727.8	67.2	910	5 BQ675260	BQ675260 AGENCOURT
5	720.6	66.5	957	5 BQ675006	BQ675006 AGENCOURT
6	689	63.6	1265	4 AF176779	AF176779 Homo sapi
7	670.6	61.9	2634	10 AY411617	AY411617 Mus muscu
8	669.6	61.8	3143	4 AK016546	AK016546 Mus muscu
9	669.6	61.8	5517	4 AK029788	AK029788 Mus muscu
10	665.2	61.4	968	5 BQ671635	BQ671635 AGENCOURT
11	662.2	61.1	1041	2 BI084090	BI084090 602869453
12	659.8	60.9	726	7 CN427231	CN427231 170004554
13	642	59.3	1292	4 AF175225	AF175225 Homo sapi
14	619.8	57.2	674	2 BE888665	BE888665 601513061
15	616	56.9	639	2 BG165205	BG165205 602344074
16	607.8	56.1	658	2 BE388695	BE388695 601286345
17	606.6	56.0	753	1 AL040398	AL040398 DKFZp434A
18	596.4	55.1	746	5 BY751026	BY751026 BY751026
19	591.2	54.6	1278	4 AF173382	AF173382 Homo sapi
20	590.6	54.5	766	6 CD803232	CD803232 UI-M-GV0-
21	579.6	53.5	748	6 CD803197	CD803197 UI-M-GV0-
22	578.6	53.4	684	6 CF533690	CF533690 UI-M-FY0-

23	575.2	53.1	729	6	CF726752	CF726752 UI-M-HB0-
24	570.6	52.7	1090	3	BM910838	BM910838 AGENCOURT
25	568.2	52.5	671	8	CX787143	CX787143 HESC3 81
26	565.2	52.2	1013	7	CN647232	CN647232 ILLUMIGEN
27	565.2	52.2	1027	7	CN646612	CN646612 ILLUMIGEN
28	562.2	51.9	683	7	CN530337	CN530337 UI-M-H00-
29	543.8	50.2	853	2	BF694553	BF694553 602081767
30	539.2	49.8	703	6	CF745813	CF745813 UI-M-GV0-
31	535.4	49.4	789	2	BG709118	BG709118 602675382
32	521.4	48.1	695	2	BG715866	BG715866 602675282
33	493.8	45.6	788	2	BG163397	BG163397 602338360
34	492.8	45.5	643	7	CN529315	CN529315 UI-M-H00-
35	486.6	44.9	853	2	BI223892	BI223892 602941035
36	480.8	44.4	897	5	BU191638	BU191638 AGENCOURT
37	470.2	43.4	831	2	BG974324	BG974324 602844075
38	454.4	42.0	1082	5	BU239476	BU239476 603322338
39	450.4	41.6	705	7	CO429380	CO429380 UI-M-HW0-
40	434.2	40.1	2292	4	CR609512	CR609512 full-length
41	426.4	39.4	663	7	CN535698	CN535698 UI-M-HS0-
42	409.8	37.8	837	8	CX501844	CX501844 JGI_XZG39
43	409.4	37.8	479	2	BI135177	BI135177 UI-M-BH3-
44	397.4	36.7	738	7	CN427234	CN427234 170006000
45	395.6	36.5	875	8	DN068337	DN068337 JGI_CABD4

ALIGNMENTS

RESULT 1	AY411615	LOCUS	DEFINITION	2649 bp	DNA	linear	GSS 16-DEC-2003
AY411615	Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.	AY411615	Homo sapiens DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.	2649 bp	DNA	linear	GSS 16-DEC-2003
ACCESSION	AY411615	GI:39767583	AY411615.1	GI:39767583			
VERSION	GSS.						
KEYWORDS	Homo sapiens (human)						
SOURCE	Homo sapiens						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE	1 (bases 1 to 2649)						
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.						
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios						
JOURNAL	Science 302 (5652), 1960-1963 (2003)						
PUBMED	14671302						
REFERENCE	2 (bases 1 to 2649)						
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.						
TITLE	Direct Submission						
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA						
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.						
FEATURES	Location/Qualifiers						
source	1. 2649						
gene	/organism="Homo sapiens"						
	/mol_type="genomic DNA"						
	/db_xref="taxon:9606"						
	<1..>2649						
	/gene="DPP8"						
	/locus_tag="HCM4283"						
ORIGIN							

Query Match 73.4%; Score 794.6; DB 10; Length 2649;  
Best Local Similarity 90.0%; Pred. No. 1e-214;  
Matches 800; Conservative 0; Mismatches 89; Indels 0; Gaps 0;



```
QY 195 GGAGGGCAGATTGATTCGGTTATCTCTAAACAGGTACAGCAAACTCTAAAGTCACCTTTTA 254
Db 1442 GCGTCCGCTCTCCAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1501
QY 255 AGATGTCGAAATAATGATTGATGCTCGAAGGAAGGATCATAGTTGATGAAGTCAGAAGGC 314
Db 1502 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1561
QY 315 TGGTATATTTTGAAGGCACCAAGACATCCCTTTTAGAGCATCACCTGTACGTAGTCAGTT 374
Db 1562 TGGTATATTTTGAAGGCACCAAGACATCCCTTTTAGAGCATCACCTGTAGTCAGTT 1621
QY 375 ACGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATCTTTGCTGCA 434
Db 1622 ACGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATCTTTGCTGCA 1681
QY 435 TCAGTCAAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGTG 494
Db 1682 TCAGTCAAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGTG 1741
QY 495 TGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACCAAGCAAGATTTT 554
Db 1742 TGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACCAAGCAAGATTTT 1801
QY 555 GGGCCACATTTTGGATTTCAGCAGGTCCTCTCTGACTATACTCTCCAGAAATTTTCT 614
Db 1802 GGGCCACATTTTGGATTTCAGCAGGTCCTCTCTGACTATACTCTCCAGAAATTTTCT 1861
QY 615 CTTTGTGAAGTACTGCTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTAC 674
Db 1862 CTTTGTGAAGTACTGCTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTAC 1921
QY 675 AGCCTGGAAGAATAATCTTACTGTGCTGTTCATATATATGTTGGTCTCAGGTGAGTTGG 734
Db 1922 AGCCTGGAAGAATAATCTTACTGTGCTGTTCATATATATGTTGGTCTCAGGTGAGTTGG 1981
QY 735 TGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGTT 794
Db 1982 TGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGTT 2041
QY 795 ATGTGGTTGATGATAGACAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCG 854
Db 2042 ATGTGGTTGATGATAGACAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCG 2101
QY 855 CTTTAAATATAAAATGGGTCAAATAGAAATGAAGATCAGGTGGAAGGACTCCAATATC 2161
Db 2102 CTTTAAATATAAAATGGGTCAAATAGAAATGAAGATCAGGTGGAAGGACTCCAATATC 2161
QY 915 TAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGATCCAGCGCTGTCCTATG 974
Db 2162 TAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGATCCAGCGCTGTCCTATG 2221
QY 975 GAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATCTTCAAGGTTGCTATTG 1034
Db 2222 GAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATCTTCAAGGTTGCTATTG 2281
QY 1035 CTGGGCCCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db 2282 CTGGGCCCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGA 2330

RESULT 2
LOCUS AY411616 2649 bp DNA linear GSS 16-DEC-2003
DEFINITION Pan troglodytes DPP8 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411616
VERSION AY411616.1 GI:39767584
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Pan.
1 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment
Location/Qualifiers
source 1..2649
organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<!.>2649
/gene="DPP8"
/locus_tag="HCM4283"

Query Match 72.9%; Score 789.8; DB 10; Length 2649;
Best Local Similarity 89.7%; Pred. No. 2.4e-213;
Matches 797; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 195 GGAGGGCAGATTTCATCCGTTATCTCTAAACAGGTACAGCAAACTCTAAAGTCACCTTTTA 254
Db 1442 GCGTCCGCTCTCCAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1501
QY 255 AGATGTCGAAATAATGATTGATGCTCGAAGGAAGGATCATAGTTGATGAAGTCAGAAGGC 314
Db 1502 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1561
QY 315 TGGTATATTTTGAAGGCACCAAGACATCCCTTTTAGAGCATCACCTGTAGTCAGTT 374
Db 1562 TGGTATATTTTGAAGGCACCAAGACATCCCTTTTAGAGCATCACCTGTAGTCAGTT 1621
QY 375 ACGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATCTTTGCTGCA 434
Db 1622 ACGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATCTTTGCTGCA 494
QY 435 TCAGTCAAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGTG 1741
Db 1682 TCAGTCAAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGTG 1741
QY 495 TGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACCAAGCAAGATTTT 554
Db 1742 TGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACCAAGCAAGATTTT 1801
QY 555 GGGCCACATTTTGGATTTCAGCAGGTCCTCTCTGACTATACTCTCCAGAAATTTTCT 614
Db 1802 GGGCCACATTTTGGATTTCAGCAGGTCCTCTCTGACTATACTCTCCAGAAATTTTCT 1861
QY 615 CTTTGTGAAGTACTGCTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTAC 674
Db 1862 CTTTGTGAAGTACTGCTGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTAC 1921
QY 675 AGCCTGGAAGAATAATCTTACTGTGCTGTTCATATATATGTTGGTCTCAGGTGAGTTGG 734
Db 1922 AGCCTGGAAGAATAATCTTACTGTGCTGTTCATATATATGTTGGTCTCAGGTGAGTTGG 1981
QY 735 TGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGTT 794
Db 1982 TGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGTT 2041
QY 795 ATGTGGTTGATGATAGACAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCG 854
Db 2042 ATGTGGTTGATGATAGACAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGCG 2101
QY 855 CTTTAAATATAAAATGGGTCAAATAGAAATGAAGATCAGGTGGAAGGACTCCAATATC 2161
Db 2102 CTTTAAATATAAAATGGGTCAAATAGAAATGAAGATCAGGTGGAAGGACTCCAATATC 2161
QY 915 TAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGATCCAGCGCTGTCCTATG 974
Db 2162 TAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGATCCAGCGCTGTCCTATG 2221
QY 975 GAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATCTTCAAGGTTGCTATTG 1034
Db 2222 GAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTTCAGATCTTCAAGGTTGCTATTG 2281
QY 1035 CTGGGCCCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db 2282 CTGGGCCCCAGTCATCTGTGGATCTTCTATGATACAGGATACACGGA 2330
```



TITLE JOURNAL COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2456 row: 1 column: 07 High quality sequence stop: 618. Location/Qualifiers		Db	Qy	GGGATCCTGTACCCGAGGCTTAAATTTGAAGCGCTTAAATATAAAATGGGTCAAT 705 AGAAATTCAGCATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATTCATTGA 939 706 AGAAATTCAGCATCAGGTGGAAGACTCCAATATCTAGCTTCTCGATATGATTCATTGA 765 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCA-T 998 766 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAAT 825 999 TAATGCAGAGTGCAGA-TATCTTTCAGGGTGTCTATGCTGGGGCCCCCAGTCACTCTGTGG 1057 826 TAATGCAGAGTGCAGATTATCTTCAAGGTTGCTATTTGCTGGGGCCCCCAGTCAATTTCTGG 885 1058 ATCTTCTAT 1066 886 GGGATCTTT 894	
	RESULT 5 BQ675006 LOCUS DEFINITION BQ675006 5', mRNA sequence. ACCESSION BQ675006.1 GI:21785840 VERSION BQ675006.1 KEYWORDS EST. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Cnatiata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 957) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2456 row: h column/ 15 High quality sequence stop: 4 High quality sequence stop: 624. Location/Qualifiers					
FEATURES source	1. 957 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6274622" /issue_type="epidermoid carcinoma, cell line" /lab_hosts="DH10B (phage-resistant)" /clone_libs="NIH_MGC_102" /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."					
	Query Match 67.2%; Score 727.8; DB 5; Length 910; Best Local Similarity 96.7%; Pred. No. 8e-196; Matches 763; Conservative 0; Mismatches 24; Indels 2; Gaps 2; 280 TGAAGGAAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339 106 TGGATCTAAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 165 340 CTCCTCTTAGAGCATCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACCAAG 399 166 CTCCTCTTAGAGCATCACCTGTACGTAGTACGTAAATCTCTGGAGAGGTGACCAAG 225 400 GCTGACTGACCGTGGCTACTCACATCTTCTGCTGATCAGTCAGCAGCTGTGACTTCTTTAT 459 226 GCTGACTGACCGTGGCTACTCACATCTTCTGCTGATCAGTCAGCAGCTGTGACTTCTTTAT 285 460 AAGTAAGTATAGTAACCAAGAGATCCACACTGTGTCTCCTTTTCAAGCTATCAAGTCC 519 286 AAGTAAGTATAGTAACCAAGAGATCCACACTGTGTGTCTCTTTTCAAGCTATCAAGTCC 345 520 TGAAGATGACCAACTTGCAAAACAAAGGAATTTTGGCCACCATTTTGGATTCAGCAGG 579 346 TGAAGATGACCAACTTGCAAAACAAAGGAATTTTGGCCACCATTTTGGATTCAGCAGG 405 580 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTAC 639 406 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTAC 465 640 ATTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTTGGAAAGAAATATCTTACTGT 699 466 ATTGTATGGGATGCTCTACAAGCTCATGATCTACAGCTTGGAAAGAAATATCTTACTGT 525 700 GCTGTTCATATATGCTGCTCTCAGGTGCAATTTGGTGAATAATTCGTTTAAAGAGTCAA 759 526 GCTGTTCATATATGCTGCTCTCAGGTGCAATTTGGTGAATAATTCGTTTAAAGAGTCAA 585 760 GTATTTTCGGCTTGAATACCCCTAGCCTCTCTAGGTATGTGGTTGTAGTAGACCAACAG 819 586 GTATTTTCGGCTTGAATACCCCTAGCCTCTCTAGGTATGTGGTTGTAGTAGACCAACAG 645 820 GGGATCCTGTCAACGAGGCTTAAATTTGAAGCGCCCTTTTAAATATAAAATGGGTCAAT 879					
ORIGIN		ORIGIN				

Db	119	TCGATCTAATATCCAAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCACCAAGA	178
Qy	340	CTCCCTTTAGAGCATACCTGTAGTACAGTTAGTAAATCCCTGGAGAGTGACAAAG	399
Db	179	CTCCCTTTAGAGCATACCTGTAGTACAGTTAGTAAATCCCTGGAGAGTGACAAAG	238
Qy	400	GTTGACTGACCGTGGCTACTCAATTCCTTGTCAGTCAGTCAGCAGCTGTGACTTCTTTAT	459
Db	239	GCTGACTGACCGTGGCTACTCAATTCCTTGTCAGTCAGTCAGCAGCTGTGACTTCTTTAT	298
Qy	460	AGATAAGTATAGTAACCAAGAGATCCACACTGTGTGTCCTTTACAGCTATCAAGTCC	519
Db	299	AGTAAGTATAGTAACCAAGAGATCCACACTGTGTGTCCTTTACAGCTATCAAGTCC	358
Qy	520	TGAAGATGCCCACTTCGAAACAAAGAAATTTTGGCCACCAATTTTGGATTACAGG	579
Db	359	TGAAGATGCCCACTTCGAAACAAAGAAATTTTGGCCACCAATTTTGGATTACAGG	418
Qy	580	TCCTCTTCTGACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTAC	639
Db	419	TCCTCTTCTGACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTAC	478
Qy	640	ATTGATGGGATGCTCTCAAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGT	699
Db	479	ATTGATGGGATGCTCTCAAGCCTCATGATCTACAGCTGGAAAGAAATATCTACTGT	538
Qy	700	GCTGTTTCATATATGCTGCTCAGGTCAGTGTGTAATATCGGTTTAAAGAGTCAA	759
Db	539	GCTGTTTCATATATGCTGCTCAGGTCAGTGTGTAATATCGGTTTAAAGAGTCAA	598
Qy	760	GTAATTCGCTTGAATA-CCCTAGCTCTTAGTGTATGTTGTAGTGATAGAC-AAC	817
Db	599	GTAATTCGCTTGAATA-CCCTAGCTCTTAGTGTATGTTGTAGTGATAGACAAAC	658
Qy	818	AGGGATCTCTGACCGAGGCTTAAATTTGAAGCGCC--TTTAAATATAAAATGGGTC	875
Db	659	AGGGATCTCTGACCGAGGCTTAAATTTGAAGCGCCCTTTTAAATATAAAATGGGTC	718
Qy	876	AAATAGAAATGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATGATTTCA	935
Db	719	AAATAGAAATGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATGATTTCA	778
Qy	936	TTGACTTAGATCGTGTGGGATCCACGGCTGTCTCTATGAGAGTACTCTCCCTGATGG	995
Db	779	TTGACTTAGATCGTGTGGGATCCACGGCTGTCTCTATGAGAGTACTCTCCCTGATGG	838
Qy	996	CATTAAATGAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCTCACTCTG-	1054
Db	839	CATTAAATGAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCTCACTCTG	898
Qy	1055	TGGATCTCTATGATACAGGATACACGGA	1083
Db	899	TGGATCTCTATGATACAGGATACACGGA	927
RESULT 6			
AF176779			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1 (bases 1 to 1265)			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
Unpublished			
REFERENCE			
2 (bases 1 to 1265)			
Hui, R.T., Liu, Y.Q., Wang, X.Y., Qin, B.M. and Sheng, H.			
Submitted (10-AUG-1999) Molecular Medicine Center for			
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,			
Bei Li Shi Lu, Beijing 100037, P.R. China			
Location/Qualifiers			
1..1265			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/tissue_type="aorta"			
268..741			
/codon_start=1			
/product="MSTP141"			
/protein_id="AAQ13657.1"			
/db_xref="GI:33338070"			
/translation="MLYKPHDLQPKYPTVLFYIGQPQVOLVNNRPFKVKYPRLLNTL			
ASLGVVYVVDNRGSGHRGLKFGAFYKMGQIEIDQVEGLQYLASRYDFDLDVRG			
IHGWSYGVGLSLMALMQRSDIPRVALIAGAPVTLWIFDYDTGTRYMGHPDQNEQ"			
ORIGIN			
Query Match 63.6%; Score 689; DB 4; Length 1265;			
Best Local Similarity 99.9%; Pred. No. 1e-184;			
Matches 700; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			
Qy	384	CTGAGAGGTGACAAGG-CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAG	442
Db	1	CTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAG	60
Qy	443	CACGTGACCTCTTTTATAAGTATAGTAACTACACAGAGAAATCCACACTGTGTCCTCT	502
Db	61	CACGTGACCTCTTTTATAAGTATAGTAACTACACAGAGAAATCCACACTGTGTCCTCT	120
Qy	503	TACAAGCTATCAAGTCCTCAAGATGACCCAACTTCGAAACAAAGAAATTTTGGGCCACC	562
Db	121	TACAAGCTATCAAGTCCTCAAGATGACCCAACTTCGAAACAAAGAAATTTTGGGCCACC	180
Qy	563	ATTTTGGATTCAGCAGGTCCTCTTCTGACTACTACTCTCCAGAGAAATTTTCTTTTGA	622
Db	181	ATTTTGGATTCAGCAGGTCCTCTTCTGACTACTACTCTCCAGAGAAATTTTCTTTTGA	240
Qy	623	AGTACTACTCGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCTGGA	682
Db	241	AGTACTACTCGATTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCTGGA	300
Qy	683	AAGAAATATCTCTACTGTGCTGTTTATATGTTGGTCTCAGGTGCAGTTGTTGAATAAT	742
Db	301	AAGAAATATCTCTACTGTGCTGTTTATATGTTGGTCTCAGGTGCAGTTGTTGAATAAT	360
Qy	743	CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTT	802
Db	361	CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTT	420
Qy	803	GTAGTGATAGACACAGGGGATCCTGTCCAGCGGCTTTAAATTTGAAGGCGCTTTTAAA	862
Db	421	GTAGTGATAGACACAGGGGATCCTGTCCAGCGGCTTTAAATTTGAAGGCGCTTTTAAA	480
Qy	863	TATAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCT	922
Db	481	TATAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCT	540
Qy	923	CGATATGATTTTCTTATGATCTGTGTTGGGATCCAGGCTGCTCTATCGAGGATAC	982
Db	541	CGATATGATTTTCTTATGATCTGTGTTGGGATCCAGGCTGCTCTATCGAGGATAC	600
Qy	983	CTCTCCCTGATGGATTAATGACAGGTCAGATATCTTCAGGTTGCTATTCTCGGGGCC	1042
Db	601	CTCTCCCTGATGGATTAATGACAGGTCAGATATCTTCAGGTTGCTATTCTCGGGGCC	660
Qy	1043	CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	1083
Db	661	CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA	701

```
RESULT 7
AY411617
LOCUS
DEFINITION
MUS musculus DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION
AY411617.1 GI:39767595
VERSION
GSS
KEYWORDS
MUS musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
2 (bases 1 to 2634)
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
Source
1..2634
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<!-->2634
/gene="DPP8"
/locus_tag="HCM4283"
ORIGIN
Query Match 61.9%; Score 670.6; DB 10; Length 2634;
Best Local Similarity 81.9%; Pred. No. 2.4e-179;
Matches 718; Conservative 0; Mismatches 159; Indels 0; Gaps 0;
Qy 207 CATTCCGTTATCTTAAACAGGTACAGCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAA 266
Db 1457 CAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1516
Qy 267 TAATGATTGATCGTGAAGGAGGATCATAGTTGATGAAGTCAGAAAGCTGGTATATTG 326
Db 1517 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1576
Qy 327 AAGGCCAACAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAAATCCCTG 386
Db 1577 AAGGCCAACAAAGACTCTCTTTGGAGCATCACCTGTACGTAGTCAGTTATGCAACCCCTG 1636
Qy 387 GAGAGGTGCAAGGCTGACTGACCGTGCTACTCACATTTCTGTCATCAGTCAGCACT 446
Db 1637 GAGAAGTGGTGGGCTGACTGACCGTGCTACTCACATCTCTGCTGCTCAGCGGCAT 1696
Qy 447 GTGACTCTTTTATAGTAAGTATAGTACACGAGAGATCCACACTGCTGCTCCCTTACA 506
Db 1697 GTGACTCTTTTATAGTAAGTATAGTACACGAGAGATCCACACTGCTGCTCCCTTACA 1756
Qy 507 AGCTATCAAGTCTGAAGTACCAACCTTGCRAAACAGGAAATTTGGGACACATTT 566
Db 1757 AACTCTCAAGTCTGAGGATGACCCAGTTCTAATAACNAGAAATTTGGGACACATTT 1816
Qy 567 TGGATTACAGAGTCTCTCTCTGACTATCTCTCAGAAATTTCTTTTGAAGTA 626
Db 1817 TGGATTACAGAGTCTCTCTCTGACTATCTCTCAGAAATTTCTTTTGAAGTA 1876
Qy 627 CTACTGGATTACATTTGATGGGATGCTCTACAAAGCTCATGATCTACAGCCCTGGAAGA 686
Db 1877 CTACTGGATTACATTTGATGGGATGCTCTACAAAGCTCATGATCTACAGCCCTGGAAGA 1936
Qy 687 AATATCTCTACTGTCTCTTTCATATATGTTGGTCTCAGGTGAGTGGTGGTGAATATCGT 746
Db 1937 AATATCTCTACTGTCTCTTTCATATATGTTGGTCTCAGGTGAGTGGTGGTGAATATCGT 1996
Qy 747 TTAAGAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTGTATGTTGTAG 806
Db 1997 TTAAGAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTGTATGTTGTAG 2056
Qy 807 TGATAGACAAACAGGGATCTCTGTCACCGAGGCTTAAATTTCAAGGCCCTTTAAATATA 866
Db 2057 TGATAGACAAACAGGGATCTCTGTCACCGAGGCTTAAATTTCAAGGCCCTTTAAATATA 2116
Qy 867 AAATGGGTCAAAATAGAAATTTGACGATCAGGTGAGGAGTCCCAATATCTAGCTTCGAT 926
Db 2117 AAATGGGTCAAAATAGAAATTTGACGATCAGGTGAGGAGTCCCAATATCTAGCTTCGAT 2176
Qy 927 ATGATTTTCATTGACTTATGATCGTGGGATCCAGGCTGCTCTATGAGGATACCTCT 986
Db 2177 ATGATTTTCATTGACTTATGATCGTGGGATCCAGGCTGCTCTATGAGGATACCTCT 2236
Qy 987 CCCTGATGGCATTAATGTCAGAGGTGATATCTTCAGGGTGGTCTATGCTGGGCCCCAG 1046
Db 2237 CCCTGATGGCATTAATGTCAGAGGTGATATCTTCAGGGTGGTCTATGCTGGGCCCCAG 2296
Qy 1047 TCATCTGTGGATCTTCTATGATACAGGATACACGGA 1083
Db 2297 TCACCTGTGGATCTTCTATGATACAGGATACACGGA 2333
RESULT 8
AK016546
LOCUS
DEFINITION
MUS musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4932434F09 product:DIPEPTIDYL PEPTIDASE 8 homolog
[Homo sapiens], full insert sequence.
ACCESSION
AK016546.1 GI:12855334
VERSION
HTC; CAP trapper.
KEYWORDS
MUS musculus (house mouse)
SOURCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
10349636
REFERENCE
2
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
11042159
REFERENCE
3
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
PUBMED
11076861
```









```
QY 940 CTTAGATCGTGGGATCCAGCGCTGGTCTCTATGAGGATACCTCTCCCTGATGCATT 999
Db 4375 CTTGGATCAGTGGGATCCAGCGCTGGTCTCTATGAGGATACCTCTCCCTGATGCATT 4434
QY 1000 AATGACAGGCTCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGAT 1059
Db 4435 AATGACAGGCTCGGATATCTTTCGGGGTGGCTATTGCTGGGGCCCAAGTCACTCTGTGGAT 4494
QY 1060 CTTCTATGATACAGGATACAGGA 1083
Db 4495 CTTCTATGATACAGGATACAGGA 4518

RESULT 10
BQ671635
LOCUS BQ671635
DEFINITION AGENCOURT_8203887 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6255645
5', mRNA sequence.
ACCESSION BQ671635
VERSION BQ671635.1 GI:21782469
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 968)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Cloning by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW2407 row: a column: 22
High quality sequence stop: 630.
FEATURES
Location/Qualifiers
1..968
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6255645"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/note="Organ: salivary gland; Vector: pOT87; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 61.4%; Score 665.2; DB 5; Length 968;
Best Local Similarity 97.6%; Pred. No. 5.8e-178;
Matches 696; Conservative 0; Mismatches 14; Indels 3; Gaps 2;
QY 280 TGAAGGAGGATCATAGTTGATGAAGTCAGAGCGCTGATATTTTGAAGGCCACCAAGA 339
Db 107 TGGATCTAATATCCCAAGTTGATGAAGTCAGAGCGCTGATATTTTGAAGGCCACCAAGA 166
QY 340 CTCCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAAATCTCTGGAGGTCACAAG 399
Db 167 CTCCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAAATCTCTGGAGGTCACAAG 226
QY 400 GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTCTTTTAT 459
```

```
Db 227 GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTCTTTTAT 286
QY 460 AAGTAAGTATAGTAACACAGAGAAATCCACTGTGTGTCCTTTTACAAGCTATCAAGTCC 519
Db 287 AAGTAAGTATAGTAACACAGAGAAATCCACTGTGTGTCCTTTTACAAGCTATCAAGTCC 346
QY 520 TGAAGATGACCAACTTGCAGAAAACAAAGAAATTTTGGGGCCCAACATTTTGAATTCAGCAGG 579
Db 347 TGAAGATGACCAACTTGCAGAAAACAAAGAAATTTTGGGGCCCAACATTTTGAATTCAGCAGG 406
QY 580 TCCTCTTCCTGACTACTCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTCGAATTCAC 639
Db 407 TCCTCTTCCTGACTACTCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTCGAATTCAC 466
QY 640 ATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGT 699
Db 467 ATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGT 526
QY 700 GCTGTTTCATATATGATGCTCTCAGGTGTCAGTGGTGAATAATCGGTTTAAAGAGTCAA 759
Db 527 GCTGTTTCATATATGATGCTCTCAGGTGTCAGTGGTGAATAATCGGTTTAAAGAGTCAA 586
QY 760 GTATTTCCGCTTGAATACCCCTAGCTCTCTAGGTTATGCTGTTAGTGTATGATGACACAG 819
Db 587 GTATTTCCGCTTGAATACCCCTAGCTCTCTAGGTTATGCTGTTAGTGTATGATGACACAG 646
QY 820 GGGATCCTGTCACCGAGGGCTTAAATTTGAAGGGCGCTTTAAATATATAAAATGGTCAAAAT 879
Db 647 GGGATCCTGTCACCGAGGGCTTAAATTTGAAGGGCGCTTTANATATATAAAATGGTCAAAAT 706
QY 880 AGAAATGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTCATTTGA 939
Db 707 AGAAATGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTCATTTGA 766
QY 940 CTTAGATC-GTGTGGGCATCCACGGCTGGTCC--TATGGAGGATACCTCTCCC 989
Db 767 CTTAGATCGTGTGGGCATCCACGGCTGGCCCTATGGGAGGATACCTCTCCC 819

RESULT 11
BQ671635
LOCUS BQ671635
DEFINITION 602869453F1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013996 5',
mRNA sequence.
ACCESSION BQ671635
VERSION BQ671635.1 GI:14502420
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1041)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW1820 row: j column: 13
High quality sequence stop: 847.
FEATURES
Location/Qualifiers
1..1041
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 61.4%; Score 665.2; DB 5; Length 968;
Best Local Similarity 97.6%; Pred. No. 5.8e-178;
Matches 696; Conservative 0; Mismatches 14; Indels 3; Gaps 2;
QY 280 TGAAGGAGGATCATAGTTGATGAAGTCAGAGCGCTGATATTTTGAAGGCCACCAAGA 339
Db 107 TGGATCTAATATCCCAAGTTGATGAAGTCAGAGCGCTGATATTTTGAAGGCCACCAAGA 166
QY 340 CTCCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAAATCTCTGGAGGTCACAAG 399
Db 167 CTCCTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTAGTAAATCTCTGGAGGTCACAAG 226
QY 400 GCTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCACTGTCTTTTAT 459
```



Db	362	AAGTATTTCGGCTTGAATACCCCTAGCCCTCTCTAGGTATATGTTGGTTAGTGATAGACAAC	421	
Qy	818	AGGGGTCCTGTACCGAGGGCTTAAATTTGAAGCGCTTTAAATATAAATAAATGGGTCAA	877	
Db	422	AGGGGATCCCTGTACCGAGGGCTTAAATTTGAAGCGCTTTAAATATAAATAAATGGGTCAA	481	
Qy	878	ATAGAAATTTGACCATCAGGTGGAAGGACTCCAATATCTAGCTTTCTCGATATGATTTCAATT	937	
Db	482	ATAGAAATTTGACCATCAGGTGGAAGGACTCCAATATCTAGCTTTCTCGATATGATTTCAATT	541	
Qy	938	GACTTAGATCGTGTGGGCATCCACCGCTGGCTCTATGAGGAGATACCTCTCCCTGATGGCA	997	
Db	542	GACTTAGATCGTGTGGGCATCCACCGCTGGCTCTATGAGGAGATACCTCTCCCTGATGGCA	601	
Qy	998	TTAATG-CAGAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCCAGTCATCTGTG	1056	
Db	602	TTAATGACAGAGGTCAGATATCTTCAAGGTTGCTATTGCTGGGGCCCCAGTCATCTGTG	661	
Qy	1057	GATCTTCTATGATACAGGATACACGGA	1083	
Db	662	GATCTTCTATGATACAGGATACACGGA	688	
RESULT 13				
AF175225		1292 bp	mRNA	linear
LOCUS				HTC 01-AUG-2003
DEFINITION				Homo sapiens tissue-type aorta MSRP135 mRNA, complete cds.
ACCESSION				AF175225
VERSION				AF175225.1
KEYWORDS				HTC.
SOURCE				Homo sapiens (human)
ORGANISM				Homo sapiens
REFERENCE				
AUTHORS				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
TITLE				1 (bases 1 to 1292)
JOURNAL				Zhao, B., Xu, H.S., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B., Wang, X.Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J., Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zheng, W.Y., Teng, C.Y., Liu, Q., Yu, L.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hui, R.T. Direct Submission
FEATURES				Submitted (04-AUG-1999) Molecular Medicine Center for Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167, Bei Li Shi Lu, Beijing 100037, P.R. China
source				Location/Qualifiers
CDS				1..1292
				/organism="Homo sapiens"
				/mol_type="mRNA"
				/db_xref="taxon:9606"
				/tissue_type="aorta"
				142..1023
				/codon_start=1
				/product="MSTP135"
				/protein_id="AAQ13650.1"
				/db_xref="GI:33338056"
				/translation="MTOLAKQRNFGPPFWIQVLFITLLQKFSLLKVLDLHQMCGCS
				TSFGSTGLERNILLXLFYGGKQVQVNNRKGKVFRLNTLPLGVVXVAGINRGS
				CPRLKGFAPKFMQGIEDQVEGLQYLSKRYDFILDRVGHWSYGGYLSMAL
				MQRSDIFRAIAGAPVTLWIFDTGYTRMGHPDQNEGYLGSVMAQKFPSEPN
				RLLLHGFLENVHFAHTSILLSFLVRAGKPYDLQIYQERHSIRVPESGHEYLHL
				HVLQNLGSRIALKVI"
ORIGIN				
Query Match				59.3%; Score 642; DB 4; Length 1292;
Best Local Similarity				96.3%; Pred. No. 2.6e-171;
Matches 678; Conservative				0; Mismatches 22; Indels 4; Gaps 2;
Qy	384	CTGGAGAGGTGACAGGCTGACGACCGTGGCTACTCACATCTTCTGTCATCAGTCAGC	443	
Db	1	CTGGAGAGGTGACAGGCTGACGACCGTGGCTACTCACATCTTCTGTCATCAGTCAGC	60	
Qy	444	ACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAAGATCCACACTGTGTGTCCTTT	503	

```
source
1. .674
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:3914582"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb."

ORIGIN
Query Match 57.2%; Score 619.8; DB 2; Length 674;
Best Local Similarity 98.5%; Pred. No. 4.5e-165;
Matches 657; Conservative 0; Mismatches 7; Indels 3; Gaps 3;
QY 343 CCCTTTAGAGCATCACCCTGACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCT 402
Db 1 CCCTTTAGAGCATCACCCTGACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCT 60
QY 403 GACTGACCGTGGCTACTCACATTCCTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAG 462
Db 61 GACTGACCGTGGCTACTCACATTCCTGCTGCATCAGTCAGCACTGTGACTTCTTTATAAG 120
QY 463 TAAGTATAGTAACACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCTGA 522
Db 121 TAAGTATAGTAACACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCTGA 180
QY 523 AGATGACCCAACTTGCACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCC 582
Db 181 AGATGACCCAACTTGCACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCC 240
QY 593 TCTTCTGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTGAGTATACATTT 642
Db 241 TCTTCTGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTGAGTATACATTT 300
QY 643 GTATGGATGCTCTACAGCCTCATGATCATAGCTGGAAGAAATATCCTACTGTGCT 702
Db 301 GTATGGATGCTCTACAGCCTCATGATCATAGCTGGAAGAAATATCCTACTGTGCT 360
QY 703 GTTCATATATGGTGGTCTCAGGTGCACTGTGTGAATATCGTTTAAAGAGTCAAGTA 762
Db 361 GTTCATATATGGTGGTCTCAGGTGCACTGTGTGAATATCGTTTAAAGAGTCAAGTA 420
QY 763 TTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGG 822
Db 421 TTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGG 480
QY 823 ATCTGTACCCGAGGGCTTAAATTTGAAGCGCCTTTAAATATATAAATGGGTCAACATGA 882
Db 481 ATCTGTACCCGAGGGCTTAAATTTGAAGCGCCTTTAAATATATAAATGGGTCAACATGA 540
QY 883 AATTGACGATCAGTGGAGG-ACCTCAATATCTAGCTTCTCGATATGATTTCACTGACT 941
Db 541 AATTGACGATCAGTGGAGGAACTCCAATATCTAGCTTCTCGATATGATTTCACTGACT 600
QY 942 TAGATCTGTGGGATCCAGGCTGTGCTATGGAGGATACCTCTCCCTGATGCATTA 1001
Db 601 TAGATCTGTGGGATCCAGG- TGGTCTTATGGAGGATA-CTTCCCTGATGGCATTA 658
QY 1002 TGCAGAG 1008
Db 659 TGCAAG 665

RESULT 15
BGL65205
LOCUS 60234074F1 NIH_MGC_89 Homo sapiens cdna clone IMAGE:4453987 5',
DEFINITION mRNA sequence.
ACCESSION BGL65205
VERSION BGL65205.1 GI:12671908
```

```
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1. (bases 1 to 639)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
National Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10245 row: d column: 20
High quality sequence stop: 639.
FEATURES
Location/Qualifiers
1. 639
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:4453987"
/tissue_type="hypertrophoma, cell line"
/clone_lib="NIH_MGC_89"
/notes="Organ: Kidney; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 56.9%; Score 616; DB 2; Length 639;
Best Local Similarity 99.7%; Pred. No. 5.4e-164;
Matches 638; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 418 CTCACATTCCTGTGTCATCAGTCAGCACTGTGACTTCTTTATAAGTATAGTAAACCA 477
Db 1 CTCACATTCCTGTGTCATCAGTCAGCACTGTGACTTCTTTATAAGTATAGTAAACCA 60
QY 478 GAAGAAATCCACACTGTGTGTCCTTTACAAGCTATCAAGTCTGAAAGATGACCAACTTG 537
Db 61 GAAGAAATCCACACTGTGTGTCCTTTACAAGCTATCAAGTCTGAAAGATGACCAACTTG 120
QY 538 CAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATAC 597
Db 121 CAAACAAAGGAATTTTGGGCCACCA-TTGGATTTCAGCAGGTCTCTTCTGACTATAC 179
QY 598 TCCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTTGATGGATGCTCTA 657
Db 180 TCCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTTGATGGATGCTCTA 239
QY 658 CAGCCTCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGGTGG 717
Db 240 CAGCCTCATGATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGGTGG 299
QY 718 TCCTCAGGTGCACTGGTGAATTAATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATAC 777
Db 300 TCCTCAGGTGCACTGGTGAATTAATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATAC 359
QY 778 CCTAGCCTCTCTAGGTTATGTGTTAGTATAGACACACAGGGGATCCTGTCAACCGAGG 837
Db 360 CCTAGCCTCTCTAGGTTATGTGTTAGTATAGACACACAGGGGATCCTGTCAACCGAGG 419
QY 838 GCTTTAAATTTGAAGGGCGCTTTAAATATATAAATGGGTCAAAATAGAAAATTCAGATCAGGT 897
Db 420 GCTTTAAATTTGAAGGGCGCTTTAAATATATAAATGGGTCAAAATAGAAAATTCAGATCAGGT 479
```

Mon May 8 13:42:01 2006

Qy	898	GGAAGGACTTCCAAATATCTAGCTTCTCGATATGATTTCACTTACCTTAGATCGTGTGGGCAT	957
Db	480	GGAAGGACTTCCAAATATCTAGCTTCTCGATATGATTTCACTTACCTTAGATCGTGTGGGCAT	539
Qy	958	CCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATAT	1017
Db	540	CCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATAT	599
Qy	1018	CTTCAG-GTTTGCCTATTGCTGGGGCCCCAGTCACTCTGTG	1056
Db	600	CTTCAGCGGTGTCTATTGCTGGGGCCCCAGTCACTCTGTG	639

Search completed: May 4, 2006, 07:56:22  
Job time : 4060.93 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2006, 01:01:26 ; Search time 342.808 Seconds  
(without alignments)  
5615.673 Million cell updates/sec

Title: us-10-825-632-8  
Perfect score: 1083  
Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatacacgga 1083

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1083	100.0	1083	3	US-10-070-464-8
2	1055.8	97.5	4309	3	US-09-976-674-14
3	791.2	73.1	2797	3	US-09-976-594-1103
4	791.2	73.1	3120	3	US-10-070-464-2
5	789.6	72.9	2671	3	US-09-976-674-2
6	779.2	71.9	4829	3	US-09-976-674-12
7	497.2	45.9	4685	3	US-09-976-674-22
8	479.2	44.2	4676	3	US-09-976-674-20
9	459	42.4	1197	3	US-10-070-464-4
10	434.2	40.1	1669	3	US-10-070-464-6
11	428.2	39.5	4523	3	US-09-976-674-8
12	344.2	31.8	2617	3	US-09-976-674-4
13	344.2	31.8	4219	3	US-09-976-674-28
14	344.2	31.8	4302	3	US-09-976-674-24
15	329.8	30.5	4180	3	US-09-976-674-36
16	329.8	30.5	4263	3	US-09-976-674-32
17	311.8	28.8	4076	3	US-09-976-674-34
18	311.8	28.8	4159	3	US-09-976-674-30
19	297.4	27.5	4037	3	US-09-976-674-40
20	297.4	27.5	4120	3	US-09-976-674-38
21	198.4	18.3	823	3	US-09-280-116-171
C 22	131	12.1	480	3	US-09-270-767-684
C 23	131	12.1	480	3	US-09-270-767-15966
C 24	129.2	11.9	612	3	US-09-392-184-31

25	91.4	8.4	2411	3	US-09-976-674-26	Sequence 26, Appl
26	62.6	5.8	2946	3	US-09-902-540-9675	Sequence 9675, Ap
c 27	62.6	5.8	14555	3	US-09-902-540-1096	Sequence 1096, Ap
28	61.2	5.7	2199	3	US-09-518-550-38	Sequence 38, Appl
29	59.6	5.5	3407	3	US-10-002-593-5	Sequence 5, Appli
30	59.6	5.5	3407	3	US-09-949-016-275	Sequence 275, App
31	59.6	5.5	3407	3	US-10-423-714-5	Sequence 5, Appli
32	59.6	5.5	3419	3	US-09-949-016-4579	Sequence 4579, Ap
33	58.2	5.4	657	3	US-09-221-0178-646	Sequence 646, App
c 34	58	5.4	2924	3	US-09-023-655-1026	Sequence 1026, Ap
35	58	5.4	2924	6	PCT-US93-07923-1	Sequence 1, Appli
36	57.8	5.3	2214	3	US-09-902-540-4124	Sequence 4124, Ap
c 37	57.8	5.3	23091	3	US-09-902-540-1204	Sequence 1204, Ap
38	55	5.1	1023	3	US-09-270-767-13726	Sequence 13726, A
39	52.4	4.8	535	3	US-09-389-681-428	Sequence 428, App
40	52.4	4.8	535	3	US-09-620-4058-428	Sequence 428, App
41	52.4	4.8	535	3	US-09-433-826B-428	Sequence 428, App
42	52.4	4.8	535	3	US-09-604-287A-428	Sequence 428, App
43	52.4	4.8	535	3	US-09-834-759-428	Sequence 428, App
44	52.4	4.8	535	3	US-09-590-751A-428	Sequence 428, App
45	52.4	4.8	535	3	US-09-551-621-428	Sequence 428, App

ALIGNMENTS

RESULT 1  
US-10-070-464-8  
; Sequence 8, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 1083  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-070-464-8  
Query Match 100.0%; Score 1083; DB 3; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 8.9e-310;  
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGAAGAGATGCCAGATCAGCTGAGTCCCTACCTTTGTTCTCCAGAGAAATTTGATAG 60  
DB 1 GGAAGAGATGCCAGATCAGCTGAGTCCCTACCTTTGTTCTCCAGAGAAATTTGATAG 60  
QY 61 ATATTTCTGCTATTGGTGGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCT 120  
DB 61 ATATTTCTGCTATTGGTGGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCT 120  
QY 121 TAGAATTCATATGAGAAATGATGAATCTGAGTGGGAAATTTATTCATGTTACATCCCC 180  
DB 121 TAGAATTCATATGAGAAATGATGAATCTGAGTGGGAAATTTATTCATGTTACATCCCC 180  
QY 181 TATGTTGGAACAACAGGAGGGCAGATTCATTCCGTTATCTAAACAGGTACAGCAATCC 240  
DB 181 TATGTTGGAACAACAGGAGGGCAGATTCATTCCGTTATCTAAACAGGTACAGCAATCC 240  
QY 241 TAAAGTCACTTTTAAAGATGTGAGAAATATGATGATGCTGAAGGAGGATCATGATTGA 300  
DB 241 TAAAGTCACTTTTAAAGATGTGAGAAATATGATGATGCTGAAGGAGGATCATGATTGA 300



Db 241 TAAAGTCACCTTTTAAAGATGTGAGAAATAATGATTGATGCTGAAGGAAGGATCATAGTTGA 300  
Qy 301 TGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 360  
Db 301 TGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 360  
Qy 361 GTACGTAGTCAGTTACGTAAATCTCTGAGAGGTTGACAAGGCTGACTGACCGTGGCTACTC 420  
Db 361 GTACGTAGTCAGTTACGTAAATCTCTGAGAGGTTGACAAGGCTGACTGACCGTGGCTACTC 420  
Qy 421 ACATTCCTGCTGATCAGTCAGCACTGCTGACTCTCTTTATAGTAAGTATAGTATACCAAGAA 480  
Db 421 ACATTCCTGCTGATCAGTCAGCACTGCTGACTCTCTTTATAGTAAGTATAGTATACCAAGAA 480  
Qy 481 GAATCCACACCTGCTGCTGCTTTTACAAGCTATCAAGTCTCTGAAGATGCCCACTTGGCAA 540  
Db 481 GAATCCACACCTGCTGCTGCTTTTACAAGCTATCAAGTCTCTGAAGATGCCCACTTGGCAA 540  
Qy 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGAGGTCCTCTTCCTGACTATATCTCC 600  
Db 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGAGGTCCTCTTCCTGACTATATCTCC 600  
Qy 601 TCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTCAGAGGTCCTCTTCCTGACTATATCTCC 660  
Db 601 TCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTCAGAGGTCCTCTTCCTGACTATATCTCC 660  
Qy 661 GCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTATATATGTTGATGCTGCTCC 720  
Db 661 GCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTATATATGTTGATGCTGCTCC 720  
Qy 721 TCAGGTGCACTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACCT 780  
Db 721 TCAGGTGCACTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACCT 780  
Qy 781 AGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTCAACCGAGGCT 840  
Db 781 AGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTCAACCGAGGCT 840  
Qy 841 TAAATTTGAAGGCGCTTTAAATATAAATGAGTCAATAGAAATTCAGCATCAGGTGGA 900  
Db 841 TAAATTTGAAGGCGCTTTAAATATAAATGAGTCAATAGAAATTCAGCATCAGGTGGA 900  
Qy 901 AGGACTCCATATCTAGCTTCTCGATGATGATTTCAATGATGATGATGATGATGATGATGATGAT 960  
Db 901 AGGACTCCATATCTAGCTTCTCGATGATGATTTCAATGATGATGATGATGATGATGATGATGAT 960  
Qy 961 GCGCTGCTCTATGAGGATACCTCTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db 961 GCGCTGCTCTATGAGGATACCTCTCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Qy 1021 CAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 1080  
Db 1021 CAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 1080  
Qy 1081 GGA 1083  
Db 1081 GGA 1083

RESULT 2  
US-09-976-674-14  
; Sequence 14, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117

; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 4309  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-14

Query Match 97.5%; Score 1055.8; DB 3; Length 4309;  
Best Local Similarity 99.5%; Pred. No. 1.8e-301;  
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;  
Qy 1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTTGATAG 60  
Db 942 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTTGATAG 1001  
Qy 61 ATATTCCTGGCTATTGGTGGTGTCCAAAGCTGAAACCAACTCCCAAGTGGTGGTAAATTTCT 120  
Db 1002 ATATTCCTGGCTATTGGTGGTGTCCAAAGCTGAAACCAACTCCCAAGTGGTGGTAAATTTCT 1061  
Qy 121 TAGAATTTCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCC 180  
Db 1062 TAGAATTTCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCC 1121  
Qy 181 TATGTTGGAAACAAGGAGGCGCAGATTCAATTCGTTATCTCTAAACAAGGTACAGCAATCCC 240  
Db 1122 TATGTTGGAAACAAGGAGGCGCAGATTCAATTCGTTATCTCTAAACAAGGTACAGCAATCCC 1181  
Qy 241 TAAAGTCACCTTTTAAAGATGTGAGAAATAATGATTGATGCTGAAGGAAGGATCATAGTTGA 300  
Db 1182 TAAAGTCACCTTTTAAAGATGTGAGAAATAATGATTGATGCTGAAGGAAGGATCATAGTTGA 1240  
Qy 301 TGAAGTCAGAAAGGCTGGTATATTTTCAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 360  
Db 1241 TGAAGTCAGAAAGGCTGGTATATTTTCAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 1300  
Qy 361 GTACGTAGTCAGTTACGTAAATCTCTGAGAGTGAAGGCTGACTGACCGTGGCTACTC 420  
Db 1301 GTACGTAGTCAGTTACGTAAATCTCTGAGAGTGAAGGCTGACTGACCGTGGCTACTC 1360  
Qy 421 ACATTCCTGCTGATCAGTCAGCACTGTGACTTTCTTTTAAAGTAAAGTATAGTAACCAAGAA 480  
Db 1361 ACATTCCTGCTGATCAGTCAGCACTGTGACTTTCTTTTAAAGTAAAGTATAGTAACCAAGAA 1420  
Qy 481 GAATCCACACCTGCTGCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAG 540  
Db 1421 GAATCCACACCTGCTGCTCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAG 1480  
Qy 541 AACAAAGGAATTTTGGGCCCAACCAATTTTGGATTTCAGCAGGTCCTCTTCCTGACTATCTCC 600  
Db 1481 AACAAAGGAATTTTGGGCCCAACCAATTTTGGATTTCAGCAGGTCCTCTTCCTGACTATCTCC 1540  
Qy 601 TCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTTGAATTTGATGGGATGCTCTACAA 660  
Db 1541 TCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTTGAATTTGATGGGATGCTCTACAA 1600  
Qy 661 GCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTCAATATATGTTGG - T 718  
Db 1601 GCCTCATGATCTACAGCTGGAAGAAATATCTACTGTGCTGTTTCAATATATGTTGGTCT 1660  
Qy 719 CCTCAGGTGAGTTGGTGAATTAATCGTTTAAAGGAGTCAAGTATTTCCGCTTCAATACC 778  
Db 1661 CCTCAGGTGAGTTGGTGAATTAATCGTTTAAAGGAGTCAAGTATTTCCGCTTCAATACC 1720  
Qy 779 CTAGCCTCTCTAGGTTATCTGTTGTAGTGATAGACAACAGGGATCTGTCCACCGAGGG 838  
Db 1721 CTAGCCTCTCTAGGTTATCTGTTGTAGTGATAGACAACAGGGATCTGTCCACCGAGGG 1780  
Qy 839 CTTAAATTTGAAGGCGCTTTTAAATATAAATGAGTCAATAGAAATTTGACGATCAGGTG 898  
Db 1781 CTTAAATTTGAAGGCGCTTTTAAATATAAATGAGTCAATAGAAATTTGACGATCAGGTG 1840

Qy	899	GAAGGACTCCAATATCTAGCTTCGCATATGATTCAATTGACATTAGACTGTGGGCATC	958
Db	1841	GAAGGACTCCAATATCTAGCTTCGCATATGATTCAATTGACATTAGACTGTGGGCATC	1900
Qy	959	CACGGCTGTCCTATGGAGGATACTCTCCCTCATGGCATTAAATGCAGAGGTCAGATATC	1018
Db	1901	CACGGCTGTCCTATGGAGGATACTCTCCCTCATGGCATTAAATGCAGAGGTCAGATATC	1960
Qy	1019	TTCAGGTTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATAC	1078
Db	1961	TTCAGGTTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATAC	2020
Qy	1079	ACCGA 1083	
Db	2021	ACCGA 2025	

RESULT 3  
US-09-976-594-1103  
; Sequence 1103,, Application US/09976594  
; Patent NO. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 1103  
; LENGTH: 2797  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1  
US-09-976-594-1103

Query Match	73.1%;	Score	791.2;	DB	3;	Length	2797;
Best Local Similarity	99.0%;	Pred.	No. 2.4e-223;				
Matches	796;	Conservative	0;	Mismatches	8;	Indels	0;
Gaps	0;						
Qy	280	TGAAGGAAGGATCATAGTTGATGAAGTCAGAAAGCTGTGTATATTTTGAAGGCACAAAGA	339				
Db	1411	TGGATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGTGTATATTTTGAAGGCACAAAGA	1470				
Qy	340	CTCCCTTTTAGAGCATCACCTGTATAGTACGTTCAGTTACGTAAATCTCGAGAGGTCACAAG	399				
Db	1471	CTCCCTTTTAGAGCATCACCTGTATAGTACGTTCAGTTACGTAAATCTCGAGAGGTCACAAG	1530				
Qy	400	GCTGACTGACGCTGGCTACTCACATTCCTGCTGCATCAGTCAGCACTGTGACTTCCTTTAT	459				
Db	1531	GCTGACTGACCGCTGGCTACTCACATTCCTGCTGCATCAGTCAGCACTGTGACTTCCTTTAT	1590				
Qy	460	AAGTAAGTATAGTAAACCAAGAAATCCACACTGTGTGTCCTTTTCAAGACTATCAAGTCC	519				
Db	1591	AAGTAAGTATAGTAAACCAAGAAATCCACACTGTGTGTCCTTTTCAAGACTATCAAGTCC	1650				
Qy	520	TGAAGATGACCCAACTTCGAAAACAAAGGAATTTTGGGCCACANTTTGGATTGAGCAGG	579				
Db	1651	TGAAGATGACCCAACTTCGAAAACAAAGGAATTTTGGGCCACANTTTTGGATTGAGCAGG	1710				
Qy	580	TCTCTTCTCTGACTACTCTCTCCAGAAAATTTTCTCTTTTGAAGACTACTCTGGATTATAC	639				
Db	1711	TCTCTTCTCTGACTACTCTCTCCAGAAAATTTTCTCTTTTGAAGACTACTCTGGATTATAC	1770				
Qy	640	ATTGTATGGGATGCTCTCAAGACCTCAAGATCTACAGCCTGGAAAGAAATATCTACTGT	699				
Db	1771	ATTGTATGGGATGCTCTCAAGACCTCAAGATCTACAGCCTGGAAAGAAATATCTACTGT	1830				

Qy	700	GC	GTGTT	CATATAT	AGTGTG	TCTCT	CAGGTG	GdGd	AGTTGGT	GAATAAT	ATCGGTTT	AAAGGAGTCA	759
Db	1831	GC	GTGTT	CATATAT	AGTGTG	TCTCT	CAGGTG	GdGd	AGTTGGT	GAATAAT	ATCGGTTT	AAAGGAGTCA	1890
Qy	760	GT	ATTTCCG	CTTGA	TACCT	TAG	CCCTCT	CT	TAGGCTT	ATGTGGT	TGTAGT	GATAGACAACAG	819
Db	1891	GT	ATTTCCG	CTTGA	TACCT	TAG	CCCTCT	CT	TAGGCTT	ATGTGGT	TGTAGT	GATAGACAACAG	1950
Qy	820	GG	GATCC	TGTC	ACCGAGG	GGCTT	AAATTT	TGA	AGCGG	CGCTTT	AAATAT	ATAAATGGGTC	879
Db	1951	GG	GATCC	TGTC	ACCGAGG	GGCTT	AAATTT	TGA	AGCGG	CGCTTT	AAATAT	ATAAATGGGTC	2010
Qy	880	AGA	AAATTT	GAC	ATCAT	GAGT	GGGA	AGCA	CTTCCA	ATATCT	TAG	CTTCTCG	939
Db	2011	AGA	AAATTT	GAC	ATCAT	GAGT	GGGA	AGCA	CTTCCA	ATATCT	TAG	CTTCTCG	2070
Qy	940	CTT	TAGAT	CGTGT	GGGCAT	PTCCA	CGGCTGGT	TCT	TATG	GAGGAT	ACCTCT	CCCTGATGG	999
Db	2071	CTT	TAGAT	CGTGT	GGGCAT	PTCCA	CGGCTGGT	TCT	TATG	GAGGAT	ACCTCT	CCCTGATGG	2130
Qy	1000	AAT	G	CAGAGG	T	CAGAT	CTT	CAGG	TGCT	TATTTG	CTGGG	CCCGCAG	1059
Db	2131	AAT	G	CAGAGG	T	CAGAT	CTT	CAGG	TGCT	TATTTG	CTGGG	CCCGCAG	2190
Qy	1060	CTT	CTAT	GATAC	CAGGATAC	CACGGA	1083						
Db	2191	CTT	CTAT	GATAC	CAGGATAC	CACGGA	2214						
RESULT 4													
US-10-070-464-2													
; Sequence 2, Application US/10070464													
; Patent No. 6881564													
; GENERAL INFORMATION:													
; APPLICANT: GORRELL, Mark Douglas													
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES													
; FILE REFERENCE: GH-007													
; CURRENT APPLICATION NUMBER: US/10/070,464													
; CURRENT FILING DATE: 2002-03-07													
; PRIOR APPLICATION NUMBER: PCT/AU00/01085													
; PRIOR FILING DATE: 2000-09-11													
; PRIOR FILING DATE: 2000-02-18													
; PRIOR APPLICATION NUMBER: AU PQ2762													
; PRIOR FILING DATE: 1999-09-10													
; NUMBER OF SEQ ID NOS: 8													
; SOFTWARE: FastSeq for Windows Version 4.0													
; SEQ ID NO 2													
; LENGTH: 3120													

QY 520 TGAAGATGACCACTTGCACAAACAAAGGAATTTTGGGCCACCAATTTGGATTGACGAGG 579  
Db 1980 TGAAGATGACCACTTGCACAAACAAAGGAATTTTGGGCCACCAATTTGGATTGACGAGG 2039  
QY 580 TCCCTCTCCCTGACATATACCTCTCCGAGAAATTTCTCTTTTGAAGTACTACTGGATTATAC 639  
Db 2040 TCCCTCTCCCTGACATATACCTCTCCGAGAAATTTCTCTTTTGAAGTACTACTGGATTATAC 2099  
QY 640 ATTCTATGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 699  
Db 2100 ATTCTATGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 2159  
QY 700 CTCTGTTTCAATATGCTGCTCCTCAGGTGCAATTTGTTGTAATAATCGGTTTAAAGAGTCAA 759  
Db 2160 CTCTGTTTCAATATGCTGCTCCTCAGGTGCAATTTGTTGTAATAATCGGTTTAAAGAGTCAA 2219  
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTTAGTGATAGACAACAG 819  
Db 2220 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTTAGTGATAGACAACAG 2279  
QY 820 GGGATCCTGTACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 879  
Db 2280 GGGATCCTGTACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 2339  
QY 880 AGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCATTGA 939  
Db 2340 AGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCATTGA 2399  
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAAT 999  
Db 2400 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAAT 2459  
QY 1000 AATGCAGAGTCCAGATATCTTCAAGGTTGCTATTCCTGGGGCCCGAGTCACTCTGTGGAT 1059  
Db 2460 AATGCAGAGTCCAGATATCTTCAAGGTTGCTATTCCTGGGGCCCGAGTCACTCTGTGGAT 2519  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
Db 2520 CTTCTATGATACAGGATACACGGA 2543

## RESULT 5

US-09-976-674-2

; Sequence 2, Application US/09976674

; Patent No. 6844180

; GENERAL INFORMATION:

; APPLICANT: Qi, Steve

; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669

; CURRENT APPLICATION NUMBER: US/09/976,674

; CURRENT FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: US 60/240,117

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 2671

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-976-674-2

Query Match 72.9%; Score 789.6; DB 3; Length 2671;  
Best Local Similarity 98.9%; Pred. No. 6.9e-223;  
Matches 795; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 280 TGAAGGAAGATCATGTTGATGAAGTCAAGGCTGGTATATTTTGAAGGCCCAAAGA 339  
Db 1534 TGGATCTATATCAAGTTGATGAAGGCTGGTATATTTTGAAGGCCCAAAGA 1593

QY 340 CTCCCTTTAGAGATCACTGTACGTAGTCACTAGTAAATCTCTGGAGAGGTGACAAG 399  
Db 1594 CTCCCTTTAGAGATCACTGTACGTAGTCACTAGTAAATCTCTGGAGAGGTGACAAG 1653  
QY 400 GCTGACTGACCGTGGCTACTCAATCTTCTGCTCATCAGTCAAGCACTGTGACTTCTTTAT 459  
Db 1654 GCTGACTGACCGTGGCTACTCAATCTTCTGCTCATCAGTCAAGCACTGTGACTTCTTTAT 1713  
QY 460 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 519  
Db 1714 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 1773  
QY 520 TGAAGATGACCACTTGCACAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACGAGG 579  
Db 1774 TGAAGATGACCACTTGCACAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACGAGG 1833  
QY 580 TCCTCTTCTGACTATACCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTATAC 639  
Db 1834 TCCTCTTCTGACTATACCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTATAC 1893  
QY 640 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 699  
Db 1894 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 1953  
QY 700 GCTGTTTCAATATGCTGCTCCTCAGGTGCAATTTGTTGTAATAATCGGTTTAAAGAGTCAA 759  
Db 1954 GCTGTTTCAATATGCTGCTCCTCAGGTGCAATTTGTTGTAATAATCGGTTTAAAGAGTCAA 2013  
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTTAGTGATAGACAACAG 819  
Db 2014 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTTGTTAGTGATAGACAACAG 2073  
QY 820 GGGATCCTGTACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 879  
Db 2074 GGGATCCTGTACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 2133  
QY 880 AGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCATTGA 939  
Db 2134 AGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCATTGA 2193  
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAAT 999  
Db 2194 CTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAAT 2253  
QY 1000 AATGCAGAGTCCAGATATCTTCAAGGTTGCTATTCCTGGGGCCCGAGTCACTCTGTGGAT 1059  
Db 2254 AATGCAGAGTCCAGATATCTTCAAGGTTGCTATTCCTGGGGCCCGAGTCACTCTGTGGAT 2313  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
Db 2314 CTTCTATGATACAGGATACACGGA 2337

## RESULT 6

US-09-976-674-12

; Sequence 12, Application US/09976674

; Patent No. 6844180

; GENERAL INFORMATION:

; APPLICANT: Qi, Steve

; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669

; CURRENT APPLICATION NUMBER: US/09/976,674

; CURRENT FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: US 60/240,117

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 12

; LENGTH: 4829

; TYPE: DNA

```

; ORGANISM: Homo sapiens
US-09-976-674-12

Query Match      71.9%; Score 779.2; DB 3; Length 4829;
Best Local Similarity 98.8%; Pred. No. 1e-219;
Matches 796; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

QY 280 TGAAGGAAGGATCATAGTTGATGAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
DB 1740 TGGATCTTAATATCAAGTTGATCAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1799

QY 340 CTCGCCCTTTAGAGCATCAGCTGTACGTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 399
DB 1800 CTCGCCCTTTAGAGCATCAGCTGTACGTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1859

QY 400 GCTGACTGACCGTGGCTACTCAGATTTCTGCTGATCAGTCAAGTCAAGTCAAGTCAAGTCAAG 459
DB 1860 GCTGACTGACCGTGGCTACTCAGATTTCTGCTGATCAGTCAAGTCAAGTCAAGTCAAGTCAAG 1919

QY 460 AAGTAAGTATAGTAACACAGAAGATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 519
DB 1920 AAGTAAGTATAGTAACACAGAAGATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 1979

QY 520 TGAAGTACCCCAACTTGCACACAGAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGG 579
DB 1980 TGAAGTACCCCAACTTGCACACAGAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGG 2039

QY 580 TCCCTCTCTGACTATACCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGGATTTTAC 639
DB 2040 TCCCTCTCTGACTATACCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGGATTTTAC 2099

QY 640 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGT 699
DB 2100 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGT 2159

QY 700 GCTGTCATATATGGTGG--TCCCTCAGGTGCAATGCTGGAATATCGTTTAAAGGATC 757
DB 2160 GCTGTCATATATGGTGGTCTCCTCAGGTGCAATGCTGGAATATCGTTTAAAGGATC 2219

QY 758 AAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTTAGTATAGACAAAC 817
DB 2220 AAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTTAGTATAGACAAAC 2279

QY 818 AGGGGATCTGTGTCACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAA 877
DB 2280 AGGGGATCTGTGTCACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAA 2339

QY 878 ATAGAAATTCAGCATCAGGTGGAAGCACTCCAATCTAGCTTCTCGATATGATTTCAAT 937
DB 2340 ATAGAAATTCAGCATCAGGTGGAAGCACTCCAATCTAGCTTCTCGATATGATTTCAAT 2399

QY 938 GACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCA 997
DB 2400 GACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCA 2459

QY 998 TTAATGAGAGGTGAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGG 1057
DB 2460 TTAATGAGAGGTGAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGG 2519

QY 1058 ATCTTCTATGATACAGGATACACGGA 1083
DB 2520 ATCTTCTATGATACAGGATACACGGA 2545
```

RESULT 7
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Oi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis

```

; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-22

Query Match      45.9%; Score 497.2; DB 3; Length 4685;
Best Local Similarity 81.3%; Pred. No. 2.3e-136;
Matches 654; Conservative 0; Mismatches 8; Indels 142; Gaps 1;

QY 280 TGAAGGAAGGATCATAGTTGATGAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
DB 1740 TGGATCTTAATATCAAGTTGATCAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1799

QY 340 CTCGCCCTTTAGAGCATCAGCTGTACGTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 399
DB 1800 CTCGCCCTTTAGAGCATCAGCTGTACGTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1859

QY 400 GCTGACTGACCGTGGCTACTCAGATTTCTGCTGATCAGTCAAGTCAAGTCAAGTCAAGTCAAG 459
DB 1860 GCTGACTGACCGTGGCTACTCAGATTTCTGCTGATCAGTCAAGTCAAGTCAAGTCAAGTCAAG 1919

QY 460 AAGTAAGTATAGTAACACAGAAGATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 519
DB 1920 AAGTAAGTATAGTAACACAGAAGATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 1979

QY 520 TGAAGTACCCCAACTTGCACACAGAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGG 579
DB 1980 TGAAGTACCCCAACTTGCACACAGAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGG 2034

QY 580 TCCCTCTCTGACTATACCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGGATTTTAC 639
DB 2035 TCCCTCTCTGACTATACCTCTCAGAAATTTTCTCTTTGAAAGTACTACTGGATTTTAC 2034

QY 640 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGT 699
DB 2035 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGT 2034

QY 700 GCTGTTTATATATGTTGGTCTCTCAGGTGCAATGCTGGAATATCGTTTAAAGGATCAA 759
DB 2035 GCTGTTTATATATGTTGGTCTCTCAGGTGCAATGCTGGAATATCGTTTAAAGGATCAA 2077

QY 760 GTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTTAGTATAGACAAACAG 819
DB 2078 GTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGTTAGTATAGACAAACAG 2137

QY 820 GGGATCTGTGTCACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAAT 879
DB 2138 GGGATCTGTGTCACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAAT 2197

QY 880 AGAAATTCAGCATCAGGTGGAAGGACTTCCAATATCTAGCTTCTCGATATGATTTCAATGA 939
DB 2198 AGAAATTCAGCATCAGGTGGAAGGACTTCCAATATCTAGCTTCTCGATATGATTTCAATGA 2257

QY 940 CTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAT 999
DB 2258 CTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAT 2317

QY 1000 AATGACAGAGGTGAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGGAT 1059
DB 2318 AATGACAGAGGTGAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGGAT 2377

QY 1060 CTTCTTATGATACAGGATACACGGA 1083
DB 2378 CTTCTTATGATACAGGATACACGGA 2401
```

```

RESULT 8
US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20

Query Match      44.2%; Score 479.2; DB 3; Length 4676;
Best Local Similarity 80.2%; Pred. No. 4.9e-131;
Matches 645; Conservative 0; Mismatches 8; Indels 151; Gaps 1;

QY 280 TGAAGGAAGGATCATAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCCAAGA 339
DB 1740 TGGATCTAATATCCAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCCAAGA 1799
QY 340 CTCCTCTTTAGAGCATCACTGACGTAGTCACTGATTAATCTGAGAGGTGACAG 399
DB 1800 CTCCTCTTTAGAGCATCACTGACGTAGTCACTGATTAATCTGAGAGGTGACAG 1859
QY 400 GCTGACTGACCGTGGTATCATATCTTGTGATCACTGAGTCACTGACTTCTTTAT 459
DB 1860 GCTGACTGACCGTGGTATCATATCTTGTGATCACTGAGTCACTGACTTCTTTAT 1919
QY 460 AGTATAGTATAGTAACAGAGAAATCCACATCTGTGTGCTCTTTTCAAGCTATCAAGTCC 519
DB 1920 AGTATAGTATAGTAACAGAGAAATCCACATCTGTGTGCTCTTTTCAAGCTATCAAGTCC 1979
QY 520 TGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCCACTTTTGGATTACAGAG 579
DB 1980 TGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCCACTTTTGGATTACAGAG 2039
QY 580 TCCTCTTCTGACTATCTCTCAAGCCTCATGATCTACAGCCTGCAAGAAATTTTCTCTTTTGAAGTACTACTGGAATTAC 639
DB 2040 TCCTCTTCTGACTATCTCTCAAGCCTCATGATCTACAGCCTGCAAGAAATTTTCTCTTTTGAAGTACTACTGGAATTAC 2099
QY 640 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGCAAGAAATTTTCTCTTTTGAAGTACTACTGGAATTAC 699
DB 2100 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGCAAGAAATTTTCTCTTTTGAAGTACTACTGGAATTAC 2159
QY 700 GCTGTTTCATATATGGTGGTCTCAGTGCAGTTGGTGAATAATCGGTTTAAAGAGTCAA 759
DB 2160 GCTGTTTCATATATGGTGGTCTCAGTGCAGTTGGTGAATAATCGGTTTAAAGAGTCAA 2179
QY 760 GTATTTCCGTTGAATACCTAGCCTCTCTAGTTATGTGGTTGTAGTATAGCAACAG 819
DB 2180 ----- 2179
QY 820 GGGATCCTGTCAACGAGGCTTAATTTGAAGGCGCCTTTTAAATATATAAATGGTCAAT 879
DB 2180 ----- 2188
QY 880 AGAAATGAGATCAGGTGGAGGACTCCATATCTAGCTTCTCGATATGATTTCAATTGA 939
DB 2189 AGAAATGAGATCAGGTGGAGGACTCCATATCTAGCTTCTCGATATGATTTCAATTGA 2248
```

```

QY 940 CTTAGATCCTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGCATT 999
DB 2249 CTTAGATCCTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGCATT 2308
QY 1000 AATGAGAGGTGATATATCTTCAAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 1059
DB 2309 AATGAGAGGTGATATATCTTCAAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGAT 2368
QY 1060 CTTCTATGATACAGGATACACGGA 1083
DB 2369 CTTCTATGATACAGGATACACGGA 2392

RESULT 9
US-10-070-464-4
; Sequence 4, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-4

Query Match      42.4%; Score 459; DB 3; Length 1197;
Best Local Similarity 80.7%; Pred. No. 2.6e-125;
Matches 616; Conservative 0; Mismatches 0; Indels 147; Gaps 1;

QY 321 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACGTAA 380
DB 1 ATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTACGTAA 60
QY 381 ATCTGTGAGAGGTGACCAAGGCTGACTGACCTGGGCTACTCACATTCTTGTGCTCAGTC 440
DB 61 ATCTGTGAGAGGTGACCAAGGCTGACTGACCTGGGCTACTCACATTCTTGTGCTCAGTC 120
QY 441 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 500
DB 121 AGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCC 180
QY 501 TTTTACAGCTATCAAGTCTGAGATGACCCCACTTGCACCAAGAAATTTTGGGCCA 560
DB 181 TTTTACAGCTATCAAGTCTGAGATGACCCCACTTGCACCAAGAAATTTTGGGCCA 240
QY 561 CCATTTTGGATTCAGCAGGCTCTCTCTGACTATACCTCCAGAAATTTTCTTTT 620
DB 241 CCATTTTGGATTCAGCAGGCTCTCTCTGACTATACCTCCAGAAATTTTCTTTT 300
QY 621 AAGTACTACTGATTTACATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCTG 680
DB 301 AAGTACTACTGATTTACATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCTG 360
QY 681 GAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTTCCTCAGGTGCGTGTGTAATA 740
DB 361 GAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTTCCTCAGGTGCGTGTGTAATA 403
QY 741 ATCGGTTTAAAGGATCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGTTTATGTGG 800
DB 404 ----- 403
```

801 TTGTAGTATAGACAAAGGGGATCCTGTCCAGGGGCTTAAATTTGAAGGGCCCTTTA 860  
404 ----- 403  
861 ATATATAAATGGGTCAATAGAAATTCAGCATCAGGTGGGAAGGATCCCAATATCTAGCTT 920  
404 -----GGGTCAATATAGAAATTCAGCATCAGGTGGGAAGGATCCCAATATCTAGCTT 453  
921 CTCGATATGATTTTCACTTACCTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAT 980  
454 CTCGATATGATTTTCACTTACCTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGGAT 513  
981 ACTCTCCCTGATGATTAATGAGAGGTCAGATATCTTCAGGGTGTCTATGCTGGGG 1040  
514 ACTCTCCCTGATGATTAATGAGAGGTCAGATATCTTCAGGGTGTCTATGCTGGGG 573  
1041 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083  
574 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 616

RESULT 10  
US-10-070-464-6  
; Sequence 6, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1669  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-070-464-6

Query Match 40.1%; Score 434.2; DB 3; Length 1669;  
Best Local Similarity 98.2%; Pred. No. 6.3e-118;  
Matches 439; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAAGGATCATAGTTGATGAGTCAAGTCAAGAGGCTGGTATATTTGAAGGCACCAAGA 339  
DB 577 TGGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCACCAAGA 636  
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTCACTGATTAATCTCGAGAGGTGACAAG 399  
DB 637 CTCCTCTTTAGAGCATCACCTGTACGTAGTCACTGATTAATCTCGAGAGGTGACAAG 696  
QY 400 GCTGACTGACCGTGGTACTCAATTTCTGCTGCATCAGTCAAGCTGTGACTCTTTAT 459  
DB 697 GCTGACTGACCGTGGTACTCAATTTCTGCTGCATCAGTCAAGCTGTGACTCTTTAT 756  
QY 460 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCC 519  
DB 757 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCC 816  
QY 520 TGAAGATGACCCAACTTGCAGAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGG 579  
DB 817 TGAAGATGACCCAACTTGCAGAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGG 876  
QY 580 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTAC 639

877 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTAC 936  
QY 640 ATTGTATGGGATGCTCTACAGGCTCTATCTACAGCCTGGAAAGAAATATCTTACTGT 699  
DB 937 ATTGTATGGGATGCTCTACAGGCTCTATCTACAGCCTGGAAAGAAATATCTTACTGT 996  
QY 700 GCTGTTTCATATATGTTGGTCTCTCAGGT 726  
DB 997 GCTGTTTCATATATGTTGGTCTCTCAGGT 1023

RESULT 11  
US-09-976-674-8  
; Sequence 8, Application US/09976674  
; Patent No. 684180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 4523  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-976-674-8

Query Match 39.5%; Score 428.2; DB 3; Length 4523;  
Best Local Similarity 96.1%; Pred. No. 5.7e-116;  
Matches 439; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 280 TGAAGGAAGGATCATAGTTGATGAGTCAAGTCAAGAGGCTGGTATATTTGAAGGCACCAAGA 339  
DB 1740 TGGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCACCAAGA 1799  
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTCACTGATTAATCTCGAGAGGTGACAAG 399  
DB 1800 CTCCTCTTTAGAGCATCACCTGTACGTAGTCACTGATTAATCTCGAGAGGTGACAAG 1859  
QY 400 GCTGACTGACCGTGGTACTCAATTTCTGCTGCATCAGTCAAGCTGTGACTCTTTAT 459  
DB 1860 GCTGACTGACCGTGGTACTCAATTTCTGCTGCATCAGTCAAGCTGTGACTCTTTAT 1919  
QY 460 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCC 519  
DB 1920 AAGTAAGTATAGTAACCAAGAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGG 1979  
QY 520 TGAAGATGACCCAACTTGCAGAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGG 579  
DB 1980 TGAAGATGACCCAACTTGCAGAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGG 2039  
QY 580 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTAC 639  
DB 2040 TCCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTAC 2099  
QY 640 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 699  
DB 2100 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGT 2159  
QY 700 GCTGTTTCATATATGTTGGTCTCTCAGGTGCAATTTGGTG 736  
DB 2160 GCTGTTTCATATATGTTGGTCTCTTATTTGCTGGG-2196

RESULT 12  
US-09-976-674-4

```
; Sequence 4, Application US/09976674
; Patent No.: 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-4

Query Match      31.8%; Score 344.2; DB 3; Length 2617;
Best Local Similarity 64.5%; Pred. No. 38-91;
Matches 514; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

QY 287 AGGATCATAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCT 346
DB 1517 AAGATCTGGTCAATGAGGAGACCAAGCTGGTGTACTTCCAGGGCACCAAGGACGCGC 1576
QY 347 TTAGAGCATCACCTGTACGTGTCAGTTACCGTAAATCTGGAGAGGTGACAAGGCTGACT 406
DB 1577 CTGGAGCACACCTCTACGTGTGTCAGCTATGAGGCGCGCGGAGATCGTACGCTCAC 1636
QY 407 GACCGTGGCTACTCATCTTCTGTCATCAGTCAGCACTGTGACTTCTTTATAAGTAAG 466
DB 1637 ACGCCCGGGCTTCTCCCATAGCTGCTCCATGAGCGGAGCAACTTCGACATGTTTCTG 1696
QY 467 TATAGTAACCAAGAAGATCCACACTGTGTGTCTCTTTTGAAGCTATCAAGTCTCAAG 526
DB 1697 TACAGCAGCGGTGAGCAGCGCGCTGCTGCACTTCAAGCTGAGCGGCCCGCCGAGC 1756
QY 527 GACCCAACTTGCACCAAGAAATTTTGGGCCACCAATTTTGAATTCAGCAGGTCTCTTT 586
DB 1757 GACCCCTCGACACAGCAGCGCGCTTCTGGGCTAGCATGATGAGGCGAGCCAGCTGCC 1816
QY 587 CCTGACTATACCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTACATTTGAT 646
DB 1817 CCGGATTTATGTTCTCCAGAGATCTTCCATTTCCACAGCGCTCGGATGTGGGCTCTAC 1876
QY 647 GGGATGCTCTACAGCGCTCATGATCTACAGCTCGAAGAAATATCTTACTGTGCTTTC 706
DB 1877 GGCATGATCTACAAGCCCAAGCGCTTGAGCAGGAGGAAGACCCACCGCTCTCTTT 1936
QY 707 ATATATGGTGTCTCAGGTGAGTGGTGAATATCGGTTTAAAGAGTCAAGTATTTC 766
DB 1937 GTATATGGAGGCGCCAGGTGAGTGGTGAATATCTCTTCAAGGGCATCAAGTACTTG 1996
QY 767 CGGTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCC 826
DB 1997 CGGCTCAACACACTGGCTCTCTGGGTACGCGGTGGTGTGATTTACGCGCAGGGGCTCC 2056
QY 827 TGTCAACGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAAATGGGTCAAATAGAA 886
DB 2057 TGTCAACGAGGCTTCCGTTTCAAGGGGCGCTGAAAAACCAAAATGGGCGAGGTGAGATC 2116
QY 887 GAGCATCAGGTGGAGGAGCTCCAAATATCTAGCTTCTCGATATGATTTTCACTTAGAT 946
DB 2117 GAGGACCAAGGTGGAGGGCTGTCAGTTCTGGCGGAGAAAGTATGGCTTTCATCGACCTG 2176
QY 947 CGTGTGGGATCCACGGCTGCTCTATGAGGAGTACTCTCTCCCTGATGGCATTAATGAG 1006
DB 2177 CGAGTTGCTCATCCATGGCTGCTCTACGGGGCTTCTCTCTGCTCATGGGCTATCCAC 2236

; Sequence 4, Application US/09976674
; Patent No.: 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 4219
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-28

Query Match      31.8%; Score 344.2; DB 3; Length 4219;
Best Local Similarity 64.5%; Pred. No. 3-78-91;
Matches 514; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

QY 287 AGGATCATAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCT 346
DB 1873 AAGATCTGGTCAATGAGGAGACCAAGCTGGTGTACTTCCAGGGCACCAAGGACGCGC 1932
QY 347 TTAGAGCATCACCTGTGACGTAGTTCAGTTACCGTAAATCTTGAGAGGTGACAAGGCTGACT 406
DB 1933 CTGGAGCACACCTCTCTAGTGTGTCAGCTATGAGGCGCGCGGAGATCGTACGCTCAC 1992
QY 407 GACCGTGGCTACTCATCTTCTGTCATCAGTCAGCACTGTGACTTCTTTATAAGTAAG 466
DB 1993 ACGCCCGGGCTTCTCCCATAGCTGCTCCATGAGCCAGAACTTCGACATGTTTCTGTCAGCCAC 2052
QY 467 TATAGTAACCAAGAAGATCCACACTGTGTGTCTCTTTTGAAGCTATCAAGTCTCTGAAGAT 526
DB 2053 TACAGCAGCGGTGAGCAGCGCGCTGCTGCTGCACTTCAAGCTGAGCGGCCCGCAGCAG 2112
QY 527 GACCCAACTTGCACCAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTT 586
DB 2113 GACCCCTCGACACAGCAGCGCGCTTCTGGGCTAGCATGATGAGGCGAGCCAGCTGCC 2172
QY 587 CCTGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTTACATTTGAT 646
DB 2173 CCGGATTTATGTTCTCCAGAGATCTTCCATTTCCACAGCGCTCGGATGTGGGCTCTAC 2232
QY 647 GGGATGCTCTACAAGCGCTCATGATCTACGCTGGAAGAAATATCTTACTGTGCTGTTTC 706
DB 2233 GGCATGATCTACAAGCGCCAGCGCTTGAGCAGGGAAGAACCCACCGCTCTCTTT 2292
QY 707 ATATATGGTGTCTCAGGTGAGTGGTGAATATCGGTTTAAAGAGTCAAGTATTTC 766
DB 2293 GTATATGAGGCGCCAGGTCAGCTGTGATTAATCTCTTCAAGGGCATCAAGTACTTG 2352
QY 767 CGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCC 826
DB 2353 CGGCTCAACACACTGGCTCTCTGGGCTACGCGGTGGTGTGATTTACGCGCAGGGGCTCC 2412
QY 827 TGTCACGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAAATGGGTCAATAGAAATT 886
```





Qy	567	TGGATT	CAGCGG	CTCT	CTTCTG	ACTATAC	TCTCTC	CAGAAATTT	CTCTTT	TGAAGTA	626
Db	2114	TGGAGG	CAGCCAG	CTGCCCC	CGGATT	TATGTT	CCTCC	CAGAGAT	CTTCC	ATTTCC	ACACGC
Qy	627	CTACTG	GATTTT	TACATT	TGATGG	ATGCT	CTAC	AGGCT	CATGAT	CTAC	AGCCTG
Db	2174	GCTCGA	TGTG	CGGCT	CTAC	GGCAT	GATCT	CA	AGCCCC	CCAC	CGCTTG
Qy	687	AATATC	CTPAC	TGCTG	CTGTT	CATAT	TATG	GTGCT	CCTC	AGGTC	AGTTG
Db	2234	AGCACCC	CACCG	CTCT	CTTT	TGAT	TATG	AGGCCCC	CCAG	GTG	AGTTG
Qy	747	TTAAGG	AGTCA	AGTAT	TTT	CGGCT	TG	ATAC	CTT	AGGTT	ATG
Db	2294	TCAAAG	GCAT	CAAGT	CTT	CGGCT	CA	CACT	GGCT	CCCT	GGGCT
Qy	807	TGATAG	ACAAC	AGGGG	ATC	CTGT	C	AGG	GGCT	TAA	ATTG
Db	2354	TGATTG	CGGCA	GGGCT	CTCT	GT	C	AGG	GGCT	TC	GGT
Qy	867	AAATGG	GTCA	AAAT	AGAA	TTG	AC	AGT	CA	AGT	CA
Db	2414	AAATGG	GC	CAAG	GTG	AGAT	CG	AG	CA	CGG	CTG
Qy	927	ATGATT	CTAT	GACT	TAGAT	CGT	GT	G	GCAT	CC	AGG
Db	2474	ATGGCT	TCAT	CGAC	CTG	AGC	GAG	TTG	CCAT	CCAT	GGCT
Qy	987	CCCTG	ATGG	CAAT	TAAT	GC	AGG	TT	CA	AGG	TT
Db	2534	CGCTCA	TGGG	CTAA	TCC	ACA	AGC	CC	CA	GGT	TT
Qy	1047	TCAC	CTGT	GGAT	CTT	CTAT	GAT	AC	AGG	ATAC	ACG
Db	2594	TCACCG	TC	TGG	CTAC	GAC	AC	AGG	GTAC	ACT	GA

Search completed: May 3, 2006, 02:13:00  
 Job time : 347.142 secs

GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: May 4, 2006, 03:41:02 ; Search time 963.428 Seconds  
(without alignments)  
9295.694 Million cell updates/sec  
Title: US-10-825-632-8  
Perfect score: 1083  
Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatacacgga 1083  
Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 9793542 seqs, 4134689005 residues  
Total number of hits satisfying chosen parameters: 19587084  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA\_Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	1083	100.0	1083	8	US-10-825-632-8
2	1055.8	97.5	4309	3	US-09-976-674-14
3	1055.8	97.5	4309	3	US-10-982-512-14
4	791.2	73.1	2649	6	US-10-054-776-1
5	791.2	73.1	2649	6	US-10-170-789-39
6	791.2	73.1	3106	7	US-10-311-035-30
7	791.2	73.1	3120	7	US-10-415-122-5
8	791.2	73.1	3120	7	US-10-825-632-2
9	791.2	73.1	3143	6	US-10-170-789-37
10	789.6	72.9	2671	3	US-09-976-674-2
11	789.6	72.9	2671	9	US-10-982-512-2
12	779.2	71.9	4829	3	US-09-976-674-12
13	779.2	71.9	4829	3	US-10-982-512-12
14	581.8	53.7	2510	7	US-10-275-505-16
15	581.8	53.7	2510	10	US-11-140-224-16
16	497.2	45.9	4685	3	US-09-976-674-22
17	497.2	45.9	4685	9	US-10-982-512-22
18	479.2	44.2	4676	3	US-09-976-674-20
19	479.2	44.2	4676	9	US-10-982-512-20
20	459	42.4	1197	8	US-10-825-632-4
21	434.2	40.1	1669	8	US-10-825-632-6
22	434.2	40.1	2830	9	US-10-956-157-2177
23	428.2	39.5	4523	3	US-09-976-674-8

24	428.2	39.5	4523	9	US-10-982-512-8	Sequence 8, Appli
25	365.4	33.7	502	3	US-09-918-995-19585	Sequence 19585, A
26	344.2	31.8	2495	7	US-10-415-122-8	Sequence 8, Appli
27	344.2	31.8	2617	3	US-09-976-674-4	Sequence 4, Appli
28	344.2	31.8	2617	9	US-10-982-512-4	Sequence 4, Appli
29	344.2	31.8	3000	7	US-10-415-122-1	Sequence 1, Appli
30	344.2	31.8	3287	7	US-10-415-122-3	Sequence 3, Appli
31	344.2	31.8	3716	9	US-10-433-757-30	Sequence 30, Appli
32	344.2	31.8	4219	3	US-09-976-674-28	Sequence 28, Appli
33	344.2	31.8	4219	9	US-10-982-512-28	Sequence 28, Appli
34	344.2	31.8	4302	3	US-09-976-674-24	Sequence 24, Appli
35	344.2	31.8	4302	9	US-10-982-512-24	Sequence 24, Appli
36	342.6	31.6	2261	6	US-10-094-749-47	Sequence 47, Appli
37	342.6	31.6	2660	7	US-10-072-012-223	Sequence 223, App
38	342.6	31.6	2660	7	US-10-072-012-225	Sequence 225, App
39	329.8	30.5	4180	3	US-09-976-674-36	Sequence 36, Appli
40	329.8	30.5	4180	9	US-10-982-512-36	Sequence 36, Appli
41	329.8	30.5	4263	3	US-09-976-674-34	Sequence 34, Appli
42	329.8	30.5	4263	9	US-10-982-512-34	Sequence 34, Appli
43	311.8	28.8	4076	3	US-09-976-674-32	Sequence 32, Appli
44	311.8	28.8	4076	9	US-10-982-512-32	Sequence 32, Appli
45	311.8	28.8	4159	3	US-09-976-674-30	Sequence 30, Appli

ALIGNMENTS

RESULT 1  
US-10-825-632-8  
; Sequence 8, Application US/10825632  
; Publication No. US20040191826A1  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: FCSB-100-Div. 1  
; CURRENT APPLICATION NUMBER: US/10/825,632  
; CURRENT FILING DATE: 2004-04-15  
; PRIOR APPLICATION NUMBER: US 10/070,464  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 1083  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-825-632-8  
Query Match 100.0%; Score 1083; DB 8; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 2.2e-304; Indels 0; Gaps 0;  
Matches 1083; Conservative 0; Mismatches 0  
QY 1 GGAAGAAGATCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAAAGAAGAAATTTGATAG 60  
Db 1 GGAAGAAGATCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAAAGAAGAAATTTGATAG 60  
QY 61 ATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCT 120  
Db 61 ATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCT 120  
QY 121 TAGAATTCATATGAAGAAATGATGAATCTGAGGTGGGAAATTTATTCATGTTACATCCCC 180  
Db 121 TAGAATTCATATGAAGAAATGATGAATCTGAGGTGGGAAATTTATTCATGTTACATCCCC 180  
QY 181 TATGTTGGAACAAGAGGAGGGCAGATTCAATTCCTGTTATCTTAAACAGGTACAGCAATCC 240  
Db 181 TATGTTGGAACAAGAGGAGGGCAGATTCAATTCCTGTTATCTTAAACAGGTACAGCAATCC 240

```
QY 241 TAAAGTCACCTTTTAAAGATGTCAGAAATATGATTGATGCTGAAGGAAGGATCATAGTTGA 300
Db 241 TAAAGTCACCTTTTAAAGATGTCAGAAATATGATTGATGCTGAAGGAAGGATCATAGTTGA 300
QY 301 TGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACT 360
Db 301 TGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACT 360
QY 361 GTACGTAGTCAGTTTACGTAAATCTCGAGAGGTGACAAGGCTGACCTGACCGTGCTACTC 420
Db 361 GTACGTAGTCAGTTTACGTAAATCTCGAGAGGTGACAAGGCTGACCTGACCGTGCTACTC 420
QY 421 ACATCTCTGCTGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAA 480
Db 421 ACATCTCTGCTGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAA 480
QY 481 GAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAG 540
Db 481 GAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAG 540
QY 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGACTATACCTCC 600
Db 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGACTATACCTCC 600
QY 601 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 660
Db 601 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 660
QY 661 GCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCC 720
Db 661 GCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCC 720
QY 721 TCAGGTGCAAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACCT 780
Db 721 TCAGGTGCAAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACCT 780
QY 781 AGCCTCTCTAGTTATGTTGCTGTAGTATAGACACAGGGGATCCTCTCACCGAGGCT 840
Db 781 AGCCTCTCTAGTTATGTTGCTGTAGTATAGACACAGGGGATCCTCTCACCGAGGCT 840
QY 841 TAAATTTGAAGGCGCTTTAAATATAAATGGTCAATAGAAATGACATCAGTGGGA 900
Db 841 TAAATTTGAAGGCGCTTTAAATATAAATGGTCAATAGAAATGACATCAGTGGGA 900
QY 901 AGGACTCAATATCTAGCTTCTCGATATGATTTCAATGATGATGATGATGATGATGATGAT 960
Db 901 AGGACTCAATATCTAGCTTCTCGATATGATTTCAATGATGATGATGATGATGATGATGAT 960
QY 961 CGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAATTAATGCAAGGTCAGATATCTT 1020
Db 961 CGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAATTAATGCAAGGTCAGATATCTT 1020
QY 1021 CAGGTTGCTATGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 1080
Db 1021 CAGGTTGCTATGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 1080
QY 1081 GGA 1083
Db 1081 GGA 1083
```

RESULT 2  
US-09-976-674-14  
; Sequence 14, Application US/09976674  
; Patent No. US2020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669

```
/ CURRENT APPLICATION NUMBER: US/09/976,674
/ CURRENT FILING DATE: 2001-10-12
/ PRIOR APPLICATION NUMBER: US 60/240,117
/ PRIOR FILING DATE: 2000-10-12
/ NUMBER OF SEQ ID NOS: 61
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 14
/ LENGTH: 4309
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-976-674-14

Query Match 97.5%; Score 1055.8; DB 3; Length 4309;
Best Local Similarity 99.5%; Pred. No. 4.2e-296;
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAAAGATTTGATAG 60
Db 942 GGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAAAGATTTGATAG 1001
QY 61 ATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCT 120
Db 1002 ATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCT 1061
QY 121 TAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCC 180
Db 1062 TAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCC 1121
QY 181 TATGTTGAAACAAGGAGGCGAGATTCATTCGTTATCTTAAACAGAGGTACAGCAATCC 240
Db 1122 TATGTTGAAACAAGGAGGCGAGATTCATTCGTTATCTTAAACAGAGGTACAGCAATCC 1181
QY 241 TAAAGTCACCTTTTAAAGATGTCAGAAATATGATTGATGCTGAAGGAAGGATCATAGTTGA 300
Db 1182 TAAAGTCACCTTTTAAAGATGTCAGAAATATGATTGATGCTGAAGGAAGGATCATAGTTGA 1240
QY 301 TGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGACATCACT 360
Db 1241 TGAAGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGACATCACT 1300
QY 361 GTACGTAGTCAGTTTACGTAAATCTCGAGAGGTGACAAGGCTGACCTGACCGTGCTACTC 420
Db 1301 GTACGTAGTCAGTTTACGTAAATCTCGAGAGGTGACAAGGCTGACCTGACCGTGCTACTC 1360
QY 421 ACATCTCTGCTGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAA 480
Db 1361 ACATCTCTGCTGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAGTATAGTAAACAGAA 1420
QY 481 GAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAG 540
Db 1421 GAATCCACACTGTGTGCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAG 1480
QY 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGACTATACCTCC 600
Db 1481 AACAAAGGAATTTTGGGCCACCAATTTTGGATTACAGAGTCTCTTCTGACTATACCTCC 1540
QY 601 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 660
Db 1541 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAA 1600
QY 661 GCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCC 718
Db 1601 GCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCC 1660
QY 719 CCTCAGGTGCAAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACC 778
Db 1661 CCTCAGGTGCAAGTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTTGAATACC 1720
QY 779 CTAGCCTCTCTAGTTATGTTGTTAGTATAGACCAACAGGGGATCCTGTCCAGGAGG 838
Db 1721 CTAGCCTCTCTAGTTATGTTGTTAGTATAGACCAACAGGGGATCCTGTCCAGGAGG 1780
QY 839 CTTAAATTTGAAGGCGCTTTAAATATAAATGGGTCAATAGAAATGACGATCAGGTG 898
```

Db 1781 CTTAAATTTGAGGCGCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGGTG 1840  
Qy 899 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATC 958  
Db 1841 GAAAGGACTCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATC 1900  
Qy 959 CAGGCTGGTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCGAGGTCAAGATATC 1018  
Db 1901 CAGGCTGGTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCGAGGTCAAGATATC 1960  
Qy 1019 TTCAAGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGAATCTTCTATGATACAGGATAC 1078  
Db 1961 TTCAAGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGAATCTTCTATGATACAGGATAC 2020  
Qy 1079 ACGGA 1083  
Db 2021 ACGGA 2025

RESULT 3  
US-10-982-512-14  
; Sequence 14, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akineanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; CURRENT FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 4309  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-982-512-14

Query Match 97.5%; Score 1055.8; DB 9; Length 4309;  
Best Local Similarity 99.5%; Pred. No. 4.2e-296;  
Matches 1080; Conservative 0; Mismatches 2; Indels 3; Gaps 2;  
Qy 1 GGAAGAAGATGCCAGATCAGCTGGAGTGGCTACCTTTGTTCTCCAGAGAAATTTGATAG 60  
Db 942 GGAAGAAGATGCCAGATCAGCTGGAGTGGCTACCTTTGTTCTCCAGAGAAATTTGATAG 1001  
Qy 61 ATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAACAACCTCCAGTGGTGGTAAATTTCT 120  
Db 1002 ATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAACAACCTCCAGTGGTGGTAAATTTCT 1061  
Qy 121 TAGAATTTCTATATGAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCC 180  
Db 1062 TAGAATTTCTATATGAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCC 1121  
Qy 181 TATGTTGGAACAAAGAGGCGCAGATTCATTTCGGTTATCTCTAAACAGGTACAGCAAAATCC 240  
Db 1122 TATGTTGGAACAAAGAGGCGCAGATTCATTTCGGTTATCTCTAAACAGGTACAGCAAAATCC 1181  
Qy 241 TAAAGTCACTTTTAAAGATGTGAGAAATATGATGCTGAAGGAGGATCATAGTTGA 300  
Db 1182 TAAAGTCACTTTTAAAGATGTGAGAAATATGATGCTGAAGGAGGATCATAGTTGA 1240  
Qy 301 TGAAGTCAGAGGCTGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 360  
Db 1241 TGAAGTCAGAGGCTGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCT 1300

Qy 361 GTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGTACTC 420  
Db 1301 GTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGTACTC 1360  
Qy 421 ACATTTCTGTCATCAGTCAGCAGCTGTdACTTCTTTATAAGTAAGTATAGTACACAGAA 480  
Db 1361 ACATTTCTGTCATCAGTCAGCAGCTGTdACTTCTTTATAAGTAAGTATAGTACACAGAA 1420  
Qy 481 GAATCCACACTGTGTGTCCTTTTACAAAGTATCAAGTCTCTGAAGATGACCCAACTTTGCAA 540  
Db 1421 GAATCCACACTGTGTGTCCTTTTACAAAGTATCAAGTCTCTGAAGATGACCCAACTTTGCAA 1480  
Qy 541 AACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGCAGGTCTCTTCTGACTATFACTCC 600  
Db 1481 AACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGCAGGTCTCTTCTGACTATFACTCC 1540  
Qy 601 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGAATGCTCTACAA 660  
Db 1541 TCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGAATGCTCTACAA 1600  
Qy 661 GCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTTCATATATGTTGGTCT 718  
Db 1601 GCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTTCATATATGTTGGTCT 1660  
Qy 719 CCTCAGTGCAGTTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC 778  
Db 1661 CCTCAGTGCAGTTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC 1720  
Qy 779 CTAGCCTCTCTAGTGTATGTTGTTAGTAGAGACCAACAGGGGATCTGTCCACGAGGG 838  
Db 1721 CTAGCCTCTCTAGTGTATGTTGTTAGTAGAGACCAACAGGGGATCTGTCCACGAGGG 1780  
Qy 839 CTTAAATTTGAAGGCGCCTTTAAATATAAAATGAGTCAAAATAGAAATTCAGCATCAGGTG 898  
Db 1781 CTTAAATTTGAAGGCGCCTTTAAATATAAAATGAGTCAAAATAGAAATTCAGCATCAGGTG 1840  
Qy 899 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATC 958  
Db 1841 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATC 1900  
Qy 959 CACGCTGGTCTCTATGGAGGATACCTCTCCTGATGGGCAATTAATGAGAGGTCAAGATATC 1018  
Db 1901 CACGCTGGTCTCTATGGAGGATACCTCTCCTGATGGGCAATTAATGAGAGGTCAAGATATC 1960  
Qy 1019 TTCAGGCTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGAATCTCTATGATACAGGATAC 1078  
Db 1961 TTCAGGCTGCTATTTGCTGGGGCCCGAGTCACTCTGTGGAATCTCTATGATACAGGATAC 2020  
Qy 1079 ACGGA 1083  
Db 2021 ACGGA 2025

RESULT 4  
US-10-054-776-1  
; Sequence 1, Application US/10054776  
; Publication No. US20030165818A1  
; GENERAL INFORMATION:  
; APPLICANT: Mark Robert Edbrooke  
; APPLICANT: Alan Peter Lewis  
; TITLE OF INVENTION: NOVEL PROTEIN  
; FILE REFERENCE: OGI042US  
; CURRENT APPLICATION NUMBER: US/10/054,776  
; CURRENT FILING DATE: 2002-01-23  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 2649  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2649)

US-10-054-776-1

```
Query Match 73.1%; Score 791.2; DB 6; Length 2649;
Best Local Similarity 99.0%; Pred. No. 4.1e-219;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAGGATCATAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339
DB 1527 TGGATCTAATATCCAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1586

QY 340 CTCCTCTTTAGAGCATCCTCTGATAGTCAGTTCGTAATCTGAGAGGTCGACAAG 399
DB 1587 CTCCTCTTTAGAGCATCCTCTGATAGTCAGTTCGTAATCTGAGAGGTCGACAAG 1646

QY 400 GCTGATGACCGTGGCTACTCAATCTTGTGTCATGATGACGACTGTGACTTCTTTAT 459
DB 1647 GCTGATGACCGTGGCTACTCAATCTTGTGTCATGATGACGACTGTGACTTCTTTAT 1706

QY 460 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCC 519
DB 1707 AAGTAAGTATAGTAACCAAGAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCC 1766

QY 520 TGAAGATGACCCAACTTGCAGAAAGAAATTTTGGGCCCACTTTTGGATTTCAGCAGG 579
DB 1767 TGAAGATGACCCAACTTGCAGAAAGAAATTTTGGGCCCACTTTTGGATTTCAGCAGG 1826

QY 580 TCCTCTTCTGACTACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTTAC 639
DB 1827 TCCTCTTCTGACTACTACTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTTAC 1886

QY 640 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCTGCGTGAATATCGTTAAAGAGTCAA 699
DB 1887 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCTGCGTGAATATCGTTAAAGAGTCAA 1946

QY 700 GCTGTTTCATATGCTGCTCCTCAGTGCAGTGTGTAATATCGTTAAAGAGTCAA 759
DB 1847 GCTGTTTCATATGCTGCTCCTCAGTGCAGTGTGTAATATCGTTAAAGAGTCAA 2006

QY 760 GTATTTCCGTTGAATACCTAGCTCTCTAGGTTATGTGTTAGTGTAGTATGATGATGATGATGAT 819
DB 2007 GTATTTCCGTTGAATACCTAGCTCTCTAGGTTATGTGTTAGTGTAGTATGATGATGATGATGAT 2066

QY 820 GGGATCTCTGTCACGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAAT 879
DB 2067 GGGATCTCTGTCACGAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAAT 2126

QY 880 AGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGA 939
DB 2127 AGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGA 2186

QY 940 CTTAGATCGTGTGGGATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAAT 999
DB 2187 CTTAGATCGTGTGGGATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAAT 2246

QY 1000 AATGACAGGTGATGATATCTTCAAGGTTGCTATGCTGGGCGCCAGTCACTCTGTGGAT 1059
DB 2247 AATGACAGGTGATGATATCTTCAAGGTTGCTATGCTGGGCGCCAGTCACTCTGTGGAT 2306

QY 1060 CTTCTATGATACAGGATACACCGA 1083
DB 2307 CTTCTATGATACAGGATACACCGA 2330
```

RESULT 5  
US-10-170-789-39  
; Sequence 39; Application US/10170789  
; Publication No. US20030180930A1  
; GENERAL INFORMATION:  
; APPLICANT: Rachel E. Meyers  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark

```
APPLICANT: Weich, Nadine  
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,  
FILE REFERENCE: 10448-191001  
CURRENT APPLICATION NUMBER: US/10/170,789  
CURRENT FILING DATE: 2002-06-13  
PRIOR APPLICATION NUMBER: US 09/797,039  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: PCT/US01/06525  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: US 60/186,061  
PRIOR FILING DATE: 2000-02-29  
PRIOR APPLICATION NUMBER: US 09/882,166  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: PCT/US01/19269  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/212,078  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 09/934,406  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/US01/26052  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: US 60/226,740  
PRIOR FILING DATE: 2000-08-21  
PRIOR APPLICATION NUMBER: US 09/861,801  
PRIOR FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: PCT/US01/16549  
PRIOR FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: US 60/205,508  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/801,267  
PRIOR FILING DATE: 2001-03-06  
PRIOR APPLICATION NUMBER: PCT/US01/07138  
PRIOR FILING DATE: 2001-03-05  
PRIOR APPLICATION NUMBER: US 60/187,454  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/829,671  
PRIOR FILING DATE: 2001-04-10  
PRIOR APPLICATION NUMBER: PCT/US01/40483  
PRIOR FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: US 60/197,508  
PRIOR FILING DATE: 2000-04-18  
PRIOR APPLICATION NUMBER: US 09/961,721  
PRIOR FILING DATE: 2001-09-24  
PRIOR APPLICATION NUMBER: PCT/US01/29904  
PRIOR FILING DATE: 2001-09-24  
PRIOR APPLICATION NUMBER: US 60/235,023  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US 10/045,367  
PRIOR FILING DATE: 2001-11-07  
PRIOR APPLICATION NUMBER: US 60/246,561  
PRIOR FILING DATE: 2000-11-07  
PRIOR APPLICATION NUMBER: US 09/801,275  
PRIOR FILING DATE: 2001-03-06  
PRIOR APPLICATION NUMBER: PCT/US01/07074  
PRIOR FILING DATE: 2001-03-05  
PRIOR APPLICATION NUMBER: US 60/187,420  
PRIOR FILING DATE: 2000-03-07  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 39  
LENGTH: 2649  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-170-789-39
```

Query Match 73.1%; Score 791.2; DB 6; Length 2649;  
Best Local Similarity 99.0%; Pred. No. 4.1e-219;  
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 280 TGAAGGAGGATCATAGTTCATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339  
DB 1527 TGGATCTAATATCCAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1586

```
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTGGAGGTCACAAG 399
Db 1587 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTGGAGGTCACAAG 1646
QY 400 GCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAGTCAGCTGCTGCTCTTTAT 459
Db 1647 GCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAGTCAGCTGCTGCTCTTTAT 1706
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACATCTGTGTCTCTTTTACAGCTATCAAGTCC 519
Db 1707 AAGTAAGTATAGTAACCAAGAAATCCACATCTGTGTCTCTTTTACAGCTATCAAGTCC 1766
QY 520 TGAAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCCAATTTTGGATTCAGCAGG 579
Db 1767 TGAAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCCAATTTTGGATTCAGCAGG 1826
QY 580 TCCTCTTCTGCACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTAC 639
Db 1827 TCCTCTTCTGCACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTAC 1886
QY 640 ATTCTATGGGATGCTCTACAGGCTCTCATGATCTACAGCTGGAAGAAATATCTACTGT 699
Db 1887 ATTCTATGGGATGCTCTACAGGCTCTCATGATCTACAGCTGGAAGAAATATCTACTGT 1946
QY 700 GCTCTTCATATATGGTGGTCTCTCAGGTGCAATTTTGAATATCTGATTTTAAAGAGTCAA 759
Db 1947 GCTCTTCATATATGGTGGTCTCTCAGGTGCAATTTTGAATATCTGATTTTAAAGAGTCAA 2006
QY 760 GTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTATGATAGACAAG 819
Db 2007 GTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTTAGTATGATAGACAAG 2066
QY 820 GGGATCTGTCCACCGGGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAAT 879
Db 2067 GGGATCTGTCCACCGGGCTTAAATTTGAAGCGCTTTAAATATAAAATGGGTCAAAT 2126
QY 880 AGAAATGACCATCAGGTGGAAGGATCCCAATATCTAGCTTCTCGATATGATTTTATGA 939
Db 2127 AGAAATGACCATCAGGTGGAAGGATCCCAATATCTAGCTTCTCGATATGATTTTATGA 2186
QY 940 CTTAGATCGTGTGGGATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAT 999
Db 2187 CTTAGATCGTGTGGGATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCAT 2246
QY 1000 AATCAGAGGTCAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGTGGAT 1059
Db 2247 AATCAGAGGTCAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCAAGTCACTCTGTGGAT 2306
QY 1060 CTTCTATGATACAGGATACCGGA 1083
Db 2307 CTTCTATGATACAGGATACCGGA 2330

RESULT 6
US-10-311-035-30.
; Sequence 30, Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Danniell B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
```

```
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yaida
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 3106
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CB1
US-10-311-035-30

Query Match 73.18; Score 791.2; DB 7; Length 3106;
Best Local Similarity 99.08; Pred. No. 4.5e-219;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
Db 1729 TGGATCTAATATCCAAAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1788
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTGGAGGTCACAAG 399
Db 1789 CTCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAATCTGGAGGTCACAAG 1848
QY 400 GCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAGTCAGCTGCTGCTCTTTAT 459
Db 1849 GCTGACTGACCGTGGCTACTCACATCTTCTGCTGCATCAGTCAGCTGCTGCTCTTTAT 1908
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACATCTGTGTGCTCTTTTACAAGCTATCAAGTCC 519
Db 1909 AAGTAAGTATAGTAACCAAGAAATCCACATCTGTGTGCTCTTTTACAAGCTATCAAGTCC 1968
QY 520 TGAAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCCAATTTTGGATTCAGCAGG 579
Db 1969 TGAAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCCAATTTTGGATTCAGCAGG 2028
QY 580 TCCTCTTCTGCACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTAC 639
Db 2029 TCCTCTTCTGCACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTAC 2088
QY 640 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGGAAGAAATATCTACTGT 699
Db 2089 ATTGTATGGGATGCTCTACAGGCTCATGATCTACAGCTGGAAGAAATATCTACTGT 2148
QY 700 GCTGTTCATATATGGTGGTCTCTCAGGTGCAATTTTGGTGGTAAATCGGTTTAAAGGAGTCAA 759
Db 2149 GCTGTTCATATATGGTGGTCTCTCAGGTGCAATTTTGGTGGTAAATCGGTTTAAAGGAGTCAA 2208
QY 760 GTATTTCCGCTTGAATACCTTAGCTCTCTAGGTTATGTTGTTAGTATGATAGACAAG 819
Db 2209 GTATTTCCGCTTGAATACCTTAGCTCTCTAGGTTATGTTGTTAGTATGATAGACAAG 2268
QY 820 GGGATCTGTCAACCGAGGGCTTTAAATTTGAAGCGGCTTTTAAATATAAAATGGGTCAAAT 879
Db 2269 GGGATCTGTCAACCGAGGGCTTTAAATTTGAAGCGGCTTTTAAATATAAAATGGGTCAAAT 2328
```



QY 880 AGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 939  
Db 2329 AGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 2388  
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 999  
Db 2389 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 2448  
QY 1000 AATCAGAGGTTCAGATATCTTTCAGGGTTCCTATGAGGATACCTCTCCCTGATGGCAAT 1059  
Db 2449 AATCAGAGGTTCAGATATCTTTCAGGGTTCCTATGAGGATACCTCTCCCTGATGGCAAT 2508  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
Db 2509 CTTCTATGATACAGGATACACGGA 2532

## RESULT 7

US-10-415-122-5  
; Sequence 5, Application US/10415122  
; Publication No. US20040053369A1  
; GENERAL INFORMATION:  
; APPLICANT: THE UNIVERSITY OF SYDNEY  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: FP15217  
; CURRENT APPLICATION NUMBER: US/10/415,122  
; CURRENT FILING DATE: 2003-08-07  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 3120  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-415-122-5

Query Match 73.1%; Score 791.2; DB 7; Length 3120;  
Best Local Similarity 99.0%; Pred. No. 4.5e-219;  
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 280 TGAAGGAAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339  
Db 1740 TGGATCTAAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799  
QY 340 CTCCTCTTTAGAGCATCACCCTGTACGTAGTTCAGTTACGTAAATCTCGAGAGGTGACAAG 399  
Db 1800 CTCCTCTTTAGAGCATCACCCTGTACGTAGTTCAGTTACGTAAATCTCGAGAGGTGACAAG 1859  
QY 400 GCTGACTGACCGTGGCTACTCACAATCTTCTGCTGATCAGTCAAGCACTGTGACTTCTTTAT 459  
Db 1860 GCTGACTGACCGTGGCTACTCACAATCTTCTGCTGATCAGTCAAGCACTGTGACTTCTTTAT 1919  
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 519  
Db 1920 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 1979  
QY 520 TGAAGATGACCCCACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 579  
Db 1980 TGAAGATGACCCCACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 2039  
QY 580 TCCTCTTCTGACTATACCTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTCAC 639  
Db 2040 TCCTCTTCTGACTATACCTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTCAC 2099  
QY 640 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTTGAAAGAAATATCTCTACTGT 699  
Db 2100 ATTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTTGAAAGAAATATCTCTACTGT 2159  
QY 700 GCTGTTTCATATATGGTGGTCTCCTCAGTGCAGTTGGTGAATTAATCGGTTTAAAGAGTCAA 759  
Db 2160 GCTGTTTCATATATGGTGGTCTCCTCAGTGCAGTTGGTGAATTAATCGGTTTAAAGAGTCAA 2219  
QY 760 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTTATGTGTTTGTAGTGATACACAG 819

Db 2220 GTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTTATGTGTTTGTAGTGATACACAG 2279  
QY 820 GGGATCTGTCCACGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAAT 879  
Db 2280 GGGATCTGTCCACGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAAT 2339  
QY 880 AGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 939  
Db 2340 AGAAATTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 2399  
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 999  
Db 2400 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 2459  
QY 1000 AATCAGAGGTTCAGATATCTTTCAGGGTTCCTATGAGGATACCTCTCCCTGATGGCAAT 1059  
Db 2460 AATCAGAGGTTCAGATATCTTTCAGGGTTCCTATGAGGATACCTCTCCCTGATGGCAAT 2519  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
Db 2520 CTTCTATGATACAGGATACACGGA 2543

## RESULT 8

US-10-825-632-2  
; Sequence 2, Application US/10825632  
; Publication No. US20040191826A1  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: FCSB-100-Div. 1  
; CURRENT APPLICATION NUMBER: US/10/825,632  
; CURRENT FILING DATE: 2004-04-15  
; PRIOR APPLICATION NUMBER: US 10/070,464  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU P05709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU P02762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 3120  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-825-632-2

Query Match 73.1%; Score 791.2; DB 8; Length 3120;  
Best Local Similarity 99.0%; Pred. No. 4.5e-219;  
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 280 TGAAGGAAGGATCATAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 339  
Db 1740 TGGATCTAAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGA 1799  
QY 340 CTCCTCTTTAGAGCATCACCCTGTACGTAGTTCAGTTACGTAAATCTCGAGAGGTGACAAG 399  
Db 1800 CTCCTCTTTAGAGCATCACCCTGTACGTAGTTCAGTTACGTAAATCTCGAGAGGTGACAAG 1859  
QY 400 GCTGACTGACCGTGGCTACTCACAATCTTCTGCTGATCAGTCAAGCACTGTGACTTCTTTAT 459  
Db 1860 GCTGACTGACCGTGGCTACTCACAATCTTCTGCTGATCAGTCAAGCACTGTGACTTCTTTAT 1919  
QY 460 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 519  
Db 1920 AAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCC 1979  
QY 520 TGAAGATGACCCCACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGG 579



QY 640 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGT 699  
DB 2115 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGT 2174  
QY 700 CATTTCATATATGGTGTCTCTCAGGTGCAATGTTGGTGAATATCGGTTTAAAGAGTCAA 759  
DB 2175 CTTGTTCATATATGGTGTCTCTCAGGTGCAATGTTGGTGAATATCGGTTTAAAGAGTCAA 2234  
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTATGTTGGTGAATATCGGTTTAAAGAGTCAA 819  
DB 2235 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTATGTTGGTGAATATCGGTTTAAAGAGTCAA 2294  
QY 820 GGGATCTGTGCAACGAGGCTTAAATTTGAAGCGCTTTAAATATATAAAATGGGTCAAAT 879  
DB 2295 GGGATCTGTGCAACGAGGCTTAAATTTGAAGCGCTTTAAATATATAAAATGGGTCAAAT 2354  
QY 880 AGAAATTGACGATCAGGTGGAAGACTTCCAAATATCTAGCTTCTCGATATGATTTCAATTGA 939  
DB 2355 AGAAATTGACGATCAGGTGGAAGACTTCCAAATATCTAGCTTCTCGATATGATTTCAATTGA 2414  
QY 940 CTTAGATCGTGTGGGATCCACGCTGGTCTTATGCTGGGGCCCGAGTCACTCTCGTGGAT 1059  
DB 2415 CTTAGATCGTGTGGGATCCACGCTGGTCTTATGCTGGGGCCCGAGTCACTCTCGTGGAT 2474  
QY 1000 AATCAGAGGTGAGATATCTTCAAGGTGCTTATGCTGGGGCCCGAGTCACTCTCGTGGAT 1059  
DB 2475 AATCAGAGGTGAGATATCTTCAAGGTGCTTATGCTGGGGCCCGAGTCACTCTCGTGGAT 2534  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
DB 2535 CTTCTATGATACAGGATACACGGA 2558

## RESULT 10

US-09-976-674-2

; Sequence 2, Application US/09976674

; Patent No. US20020115843A1

; GENERAL INFORMATION:

; APPLICANT: Qi, Steve

; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669

; CURRENT APPLICATION NUMBER: US/09/976,674

; CURRENT FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: US 60/240,117

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 2671

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-976-674-2

Query Match 72.9%; Score 789.6; DB 3; Length 2671;  
Best Local Similarity 98.9%; Pred. No. 1.2e-218;  
Matches 795; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 280 TGAAGGAAGATCATAGTTGATGAAGTCAAGAGGCTGGTATATTTGAAGGCCACCAAGA 339  
DB 1534 TGGATCTAATATCAAGTTGATGAAGTCAAGAGGCTGGTATATTTGAAGGCCACCAAGA 1593  
QY 340 CTCCTTTAGACATCACTCTGAGTGTAGTACGTAACTCTGAGAGGTGACAAG 399  
DB 1594 CTCCTTTAGACATCACTCTGAGTGTAGTACGTAACTCTGAGAGGTGACAAG 1653  
QY 400 GCTGACTGACCGTGGTACTACATCTTGTGCTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 459  
DB 1654 GCTGACTGACCGTGGTACTACATCTTGTGCTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1713  
QY 460 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 519

DB 1714 AAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCC 1773  
QY 520 TGAAGATGACCCAACTTGCAAAAACAAAGAAATTTTGGGCCAACATTTTGGATTTCAAGCAGG 579  
DB 1774 TGAAGATGACCCAACTTGCAAAAACAAAGAAATTTTGGGCCAACATTTTGGATTTCAAGCAGG 1833  
QY 580 TCCTCTTCTGCTACTATCTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGATTTTAC 639  
DB 1834 TCCTCTTCTGCTACTATCTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGATTTTAC 1893  
QY 640 ATTGTATGGGATGCTCTTACAAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGT 699  
DB 1894 ATTGTATGGGATGCTCTTACAAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGT 1953  
QY 700 GCTGTTTCATATATATGCTGCTCTCAGGTGCAATGTTGGTGAATATCGGTTTAAAGAGTCAA 759  
DB 1954 GCTGTTTCATATATATGCTGCTCTCAGGTGCAATGTTGGTGAATATCGGTTTAAAGAGTCAA 2013  
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTATGTTGGTGAATATCGGTTTAAAGAGTCAA 819  
DB 2014 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTATGTTGGTGAATATCGGTTTAAAGAGTCAA 2073  
QY 820 GGGATCTGTGCAACGAGGCTTAAATTTGAAGCGCTTTAAATATATAAAATGGGTCAAAT 879  
DB 2074 GGGATCTGTGCAACGAGGCTTAAATTTGAAGCGCTTTAAATATATAAAATGGGTCAAAT 2133  
QY 880 AGAAATTGACGATCAGGTGGAAGACTTCCAAATATCTAGCTTCTCGATATGATTTCAATTGA 939  
DB 2134 AGAAATTGACGATCAGGTGGAAGACTTCCAAATATCTAGCTTCTCGATATGATTTCAATTGA 2193  
QY 940 CTTAGATCGTGTGGGATCCACGCTGGTCTTATGCTGGGGCCCGAGTCACTCTCGTGGAT 999  
DB 2194 CTTAGATCGTGTGGGATCCACGCTGGTCTTATGCTGGGGCCCGAGTCACTCTCGTGGAT 2253  
QY 1000 AATCAGAGGTGAGATATCTTCAAGGTGCTTATGCTGGGGCCCGAGTCACTCTCGTGGAT 1059  
DB 2254 AATCAGAGGTGAGATATCTTCAAGGTGCTTATGCTGGGGCCCGAGTCACTCTCGTGGAT 2313  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
DB 2314 CTTCTATGATACAGGATACACGGA 2337

## RESULT 11

US-10-982-512-2

; Sequence 2, Application US/10982512

; Publication No. US20050059081A1

; GENERAL INFORMATION:

; APPLICANT: Qi, Steve

; APPLICANT: Akinsanya, Karen

; APPLICANT: Riviere, Pierre

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

; FILE REFERENCE: 70669

; CURRENT APPLICATION NUMBER: US/10/982,512

; CURRENT FILING DATE: 2004-11-05

; PRIOR APPLICATION NUMBER: US/09/976,674

; PRIOR FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: US 60/240,117

; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 2671

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-982-512-2

Query Match 72.9%; Score 789.6; DB 9; Length 2671;  
Best Local Similarity 98.9%; Pred. No. 1.2e-218;  
Matches 795; Conservative 0; Mismatches 9; Indels 0; Gaps 0;



```
/ APPLICANT: Qi, Steve
/ APPLICANT: Akinsanya, Karen
/ APPLICANT: Riviere, Pierre
/ APPLICANT: Junien, Jean-Louis
/ TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
/ FILE REFERENCE: 70669
/ CURRENT APPLICATION NUMBER: US/10/982,512
/ CURRENT FILING DATE: 2004-11-05
/ PRIOR APPLICATION NUMBER: US/09/976,674
/ PRIOR FILING DATE: 2001-10-12
/ PRIOR APPLICATION NUMBER: US 60/240,117
/ PRIOR FILING DATE: 2000-10-12
/ NUMBER OF SEQ ID NOS: 61
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 12
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-982-512-12

Query Match
Best Local Similarity 71.9%; Score 779.2; DB 9; Length 4829;
Matches 796; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

QY 280 TGAAGGAAGGATCATAGTTGATGAGTCAGAGGCTGGTATATTTGAAGGCACCAAGA 339
DB 1740 TGGATCTAATATCAAGTTGATGAGTCAGAGGCTGGTATATTTGAAGGCACCAAGA 1799
QY 340 CTCCTTTTAGAGCATCACTGTACGTAGTACGTAAATCTGAGAGGTGACAAG 399
DB 1800 CTCCTTTTAGAGCATCACTGTACGTAGTACGTAAATCTGAGAGGTGACAAG 1859
QY 400 GCTGACTGACCGTGGCTACTCACAATCTTCTGTCATCAGTCAGCAGCTGACTTCTTAT 459
DB 1860 GCTGACTGACCGTGGCTACTCACAATCTTCTGTCATCAGTCAGCAGCTGACTTCTTAT 1919
QY 460 AAGTAAGTATAGTAACCAAGAGATCCACATCTGTGTCCTTTTCAAGTATCAAGTCC 519
DB 1920 AAGTAAGTATAGTAACCAAGAGATCCACATCTGTGTCCTTTTCAAGTATCAAGTCC 1979
QY 520 TGAAGATGCCCACTGCAAGCAAGAGATTTTGGCCACCAATTTTGAATTCAGCAGG 579
DB 1980 TGAAGATGCCCACTGCAAGCAAGAGATTTTGGCCACCAATTTTGAATTCAGCAGG 2039
QY 580 TCCTCTCTGACTATACCTCTCCAGAAATTTCTTTTGAAGATCTACTGGATTAC 639
DB 2040 TCCTCTCTGACTATACCTCTCCAGAAATTTCTTTTGAAGATCTACTGGATTAC 2099
QY 640 ATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTCGAAGAAATATCTACTGT 699
DB 2100 ATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTCGAAGAAATATCTACTGT 2159
QY 700 GCTGTTTCATATATGGTGG--TCTCTCAGGTGAGTTGGTGAATAATCGGTTTAAAGGAGTC 757
DB 2160 GCTGTTTCATATATGGTGGTCTCTCTCAGGTGAGTTGGTGAATAATCGGTTTAAAGGAGTC 2219
QY 758 AAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTATATGTTGTTAGTAGTAGACAAC 817
DB 2220 AAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTATATGTTGTTAGTAGACAAC 2279
QY 818 AGGGGATCTCTGTACCGAGGCTTAAATTTGAAGCGCTTTTAAATATAAAATGGGTCAA 877
DB 2280 AGGGGATCTCTGTACCGAGGCTTAAATTTGAAGCGCTTTTAAATATAAAATGGGTCAA 2339
QY 878 ATAGAAATTTGAGCATCAGGTGGAAGGATCTCAATATCTAGCTTCTCGATATGATTTCAAT 937
DB 2340 ATAGAAATTTGAGCATCAGGTGGAAGGATCTCAATATCTAGCTTCTCGATATGATTTCAAT 2399
QY 938 GACTTAGATCGTGTGGGATCCACCGCTGGTCTTATGGAGGATACCTCTCCCTGGTGGCA 997
DB 2400 GACTTAGATCGTGTGGGATCCACCGCTGGTCTTATGGAGGATACCTCTCCCTGGTGGCA 2459
QY 998 TTAATGACAGGTTCAGATATCTTACGGGTTCCTATTCGTGGGGCCCCCAGTCACTCTGTGG 1057
```

```
DB 2460 TTAATGACAGGTTCAGATATCTTACGGGTTCCTATTCGTGGGGCCCCCAGTCACTCTGTGG 2519
QY 1058 ATCTTCTATGATACAGGATACACGGA 1083
DB 2520 ATCTTCTATGATACAGGATACACGGA 2545

RESULT 14
US-10-275-505-16
; Sequence 16, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DELEGANE, Angelo M.; LAL, Preeti G.
; APPLICANT: HAFALIA, April J.A.; PATTERSON, Chandra
; APPLICANT: WALIA, Nazinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Valda; ELLIOTT, Vicki S.
; APPLICANT: NGUYEN, Dannel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Junning; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dyoung Aina M.
; APPLICANT: REDDY, Roopa M.; YUE, Henry
; APPLICANT: TANG, Y. Tom
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0085 USN
; CURRENT APPLICATION NUMBER: US/10/275,505
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US01/14651
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/209,402
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/207,477
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/205,803
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 60/203,566
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/202,082
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 2510
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 376067CB1
US-10-275-505-16

Query Match
Best Local Similarity 53.7%; Score 581.8; DB 7; Length 2510;
Matches 586; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 289 GATCATAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTT 348
DB 1410 GATCCAAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTT 1469
QY 349 AGAGCATCACTGTAGTCAGTTAGTAAATCCTGGAGGCTGACAGGCTGACTGA 408
DB 1470 AGAGCATCACTGTAGTCAGTTAGTAAATCCTGGAGGCTGACAGGCTGACTGA 1529
QY 409 CCGTGGCTACTCAGATTCTTCTGTCAGTCAGTCAGCTGTCATCTTTTAAAGTAAGTA 468
DB 1530 CCGTGGCTACTCAGATTCTTCTGTCAGTCAGTCAGCTGTCATCTTTTAAAGTAAGTA 1589
QY 469 TAGTAAACGAGAGATCCACACTGTGTCCTTTTCAAGCTATCAAGTCTCAAGATGA 528
DB 1590 TAGTAAACGAGAGATCCACACTGTGTCCTTTTCAAGCTATCAAGTCTCAAGATGA 1649
QY 529 CCCAACTTGCAGAAACAAAGGAATTTTGGGCCCAACATTTTGGATTTCAGACAGGCTCTCTTC 588
```

Db 1650 CCTAACTTGCACAAACAAAGGAATTTTGGGCCACCATTTTGGGATTCAGCAGGTCCTCTTCC 1709  
Qy 589 TGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGG 648  
Db 1710 TGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGG 1769  
Qy 649 GATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTGCTGTTTCAT 708  
Db 1770 GATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTGCTGTTTCAT 1829  
Qy 709 ATATGGTGGTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCG 768  
Db 1830 ATATGGTGGTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCG 1889  
Qy 769 CTTGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAAACAGGGGATCTCTG 828  
Db 1890 CTTGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAAACAGGGGATCTCTG 1949  
Qy 829 TCACCGAGGCTTAAATTTGAAGGGCCTTTTAAATATAAAATGGGTCAATAG 881  
Db 1950 TCACCGAGGCTTAAATTTGAAGGGCCTTTTAAATATAAAATGGGTCAATAG 2002

RESULT 15  
US-11-140-224-16  
; Sequence 16, Application US/11140224  
; Publication No. US20050227280A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: DELEANE, Angelo M.; LAL, Preeti G.  
; APPLICANT: HAFALIA, April J.A.; PATTERSON, Chandra  
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam  
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.  
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.  
; APPLICANT: AZIMZAI, Valda; ELLIOTT, Vicki S.  
; APPLICANT: NGUYEN, Dannel B.; GANDHI, Ameena R.  
; APPLICANT: YANG, Junning; HERNANDEZ, Roberto  
; APPLICANT: POLICKY, Jennifer L.; LU, Dyrung Aina M.  
; APPLICANT: REDDY, Roopa M.; YUE, Henry  
; APPLICANT: TANG, Y. Tom  
; TITLE OF INVENTION: PROTEASES  
; FILE REFERENCE: PI-0085 USN  
; CURRENT APPLICATION NUMBER: US/11/140,224  
; PRIOR FILING DATE: 2005-05-31  
; PRIOR APPLICATION NUMBER: US/10/275,505  
; PRIOR FILING DATE: 2002-11-04  
; PRIOR APPLICATION NUMBER: PCT/US01/14651  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/209,402  
; PRIOR FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: 60/207,477  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/205,803  
; PRIOR FILING DATE: 2000-05-17  
; PRIOR APPLICATION NUMBER: 60/203,566  
; PRIOR FILING DATE: 2000-05-11  
; PRIOR APPLICATION NUMBER: 60/202,082  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PERL Program  
; SEQ ID NO 16  
; LENGTH: 2510  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 376067CB1  
US-11-140-224-16

Query Match 53.7%; Score 581.8; DB 10; Length 2510;  
Best Local Similarity 98.8%; Pred. No. 4.2e-158;  
Matches 586; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 289 GATCATAGTTGATGAAGTCAGAAAGGCTGTATATTTTGAAGGCACAAAGACTCCCTTTT 348  
Db 1410 GATCCAAGTTGATGAAGTCAGAAAGGCTGTATATTTTGAAGGCACAAAGACTCCCTTTT 1469  
Qy 349 AGAGCATCACTGTGATGCTAGTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGA 408  
Db 1470 AGAGCATCACTGTGATGCTAGTACGTAAATCTCTGGAGAGGTGACAAAGGCTGACTGA 1529  
Qy 409 CCGTGGCTACTCACATTTCTGTGCTAGTCAGTCAAGCTGCTGACTCTTTTATTAAGTAAGTA 468  
Db 1530 CCGTGGCTACTCACATTTCTGTGCTAGTCAGTCAAGCTGCTGACTCTTTTATTAAGTAAGTA 1589  
Qy 469 TAGTAACACAGAAAGTCCACACTGTGTCTCTTTCACAGCTATCAAGTCTCTGAAGATGA 528  
Db 1590 TAGTAACACAGAAAGTCCACACTGTGTCTCTTTCACAGCTATCAAGTCTCTGAAGATGA 1649  
Qy 529 CCCAACTTGGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCC 588  
Db 1650 CCCAACTTGGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCC 1709  
Qy 589 TGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGG 648  
Db 1710 TGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTCATTTGATGG 1769  
Qy 649 GATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTGCTGTTTCAT 708  
Db 1770 GATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTGCTGTTTCAT 1829  
Qy 709 ATATGGTGGTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCG 768  
Db 1830 ATATGGTGGTCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCG 1889  
Qy 769 CTTGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAAACAGGGGATCTCTG 828  
Db 1890 CTTGAATACCTTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAAACAGGGGATCTCTG 1949  
Qy 829 TCACCGAGGCTTAAATTTGAAGGGCCTTTTAAATATAAAATGGGTCAATAG 881  
Db 1950 TCACCGAGGCTTAAATTTGAAGGGCCTTTTAAATATAAAATGGGTCAATAG 2002

Search completed: May 4, 2006, 11:21:59  
Job time : 964.428 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2006, 03:49:39 ; Search time 1479.29 Seconds  
(without alignments)  
2981.544 Million cell updates/sec

Title: US-10-825-632-8  
Perfect score: 1083  
Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatacacgga 1083

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9306428 seqs, 2036268586 residues

Total number of hits satisfying chosen parameters: 18612856

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.New.\*  
1: /SID55/ptodata/2/pubpna/US08\_NEW\_PUB.seq1.\*  
2: /SID55/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /SID55/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /SID55/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
5: /SID55/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
6: /SID55/ptodata/2/pubpna/US09\_NEW\_PUB.seq1.\*  
7: /SID55/ptodata/2/pubpna/US09\_NEW\_PUB.seq1.\*  
8: /SID55/ptodata/2/pubpna/US09\_NEW\_PUB.seq2.\*  
9: /SID55/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
10: /SID55/ptodata/2/pubpna/US10\_NEW\_PUB.seq1.\*  
11: /SID55/ptodata/2/pubpna/US10\_NEW\_PUB.seq2.\*  
12: /SID55/ptodata/2/pubpna/US10\_NEW\_PUB.seq3.\*  
13: /SID55/ptodata/2/pubpna/US10\_NEW\_PUB.seq4.\*  
14: /SID55/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
15: /SID55/ptodata/2/pubpna/US11\_NEW\_PUB.seq2.\*  
16: /SID55/ptodata/2/pubpna/US11\_NEW\_PUB.seq3.\*  
17: /SID55/ptodata/2/pubpna/US11\_NEW\_PUB.seq4.\*  
18: /SID55/ptodata/2/pubpna/US11\_NEW\_PUB.seq5.\*  
19: /SID55/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	791.2	73.1	2649	17	US-11-151-601-21
2	791.2	73.1	3143	17	US-11-151-601-19
3	118	10.9	1346	7	US-09-925-065A-669313
4	68.	6.3	2238	18	US-11-079-463-1186
5	59.6	5.5	2217	18	US-11-208-288-3
6	59.6	5.5	2301	10	US-10-522-789-1
7	59.6	5.5	3332	18	US-11-208-288-1
8	59.6	5.5	3407	11	US-10-501-035-34
9	58.4	5.4	4852	17	US-11-136-527-2130
10	54.8	5.1	1884	18	US-11-079-463-2336
11	54.8	5.1	2283	18	US-11-208-288-5
12	52.4	4.8	535	18	US-11-226-869-428
13	52.4	4.8	2788	9	US-10-505-928-476
14	52.4	4.8	2814	17	US-11-186-284-54

Sequence 168, App  
Sequence 54125, A  
Sequence 54125, A  
Sequence 7940, Ap  
Sequence 621349, A  
Sequence 13487, A  
Sequence 28205, A  
Sequence 129442, A  
Sequence 742851, A  
Sequence 92266, A  
Sequence 193508, A  
Sequence 806917, A  
Sequence 37, Appl  
Sequence 83687, A  
Sequence 184927, A  
Sequence 798336, A  
Sequence 515055, A  
Sequence 1128454, A  
Sequence 457124, A  
Sequence 797097, A  
Sequence 797098, A  
Sequence 718581, A  
Sequence 718582, A  
Sequence 67861, A  
Sequence 401937, A  
Sequence 469031, A  
Sequence 1082440, A  
Sequence 13396, A  
Sequence 546889, A  
Sequence 524839, A

## ALIGNMENTS

## RESULT 1

US-11-151-601-21  
; Sequence 21, Application US/11151601  
; Publication No. US20060003413A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Welch, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,  
; FILE REFERENCE: AND PROTEASE FAMILY MEMBERS AND USES THEREOF  
; FILE REFERENCE: MPI00-054P1RCP10MIDVIM  
; CURRENT APPLICATION NUMBER: US/11/151.601  
; CURRENT FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: US 10/170,789  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/882,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/212,078  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/934,406  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US01/26052  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,740  
; PRIOR FILING DATE: 2000-08-21

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-151-601-21

Query Match      73.1%; Score 791.2; DB 17; Length 2649;
Best Local Similarity 99.0%; Pred. No. 5.7e-208;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAGGATCATAGTTCATGATCAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
Db      |||
QY 1527 TGGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1586
Db      |||
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCGAGAGGTCACAAAG 399
Db      |||
QY 1587 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCGAGAGGTCACAAAG 1646
Db      |||
QY 400 GCTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGTCAGCTGCTGACTTCTTTAT 459
Db      |||
QY 1647 GCTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGTCAGCTGCTGACTTCTTTAT 1706
Db      |||
QY 460 AAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 519
Db      |||
QY 1707 AAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTACAGCTATCAAGTCC 1766
Db      |||
QY 520 TGRAGATGACCCACTTGCARACAAAGAAATTTGGGCCACCATTTTGGATTCAGCAGG 579
Db      |||
QY 1767 TGRAGATGACCCACTTGCARACAAAGAAATTTGGGCCACCATTTTGGATTCAGCAGG 1826
Db      |||
QY 580 TCCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTAC 639
Db      |||
QY 1827 TCCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTAC 1886
Db      |||
QY 640 ATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGT 699
Db      |||
QY 1887 ATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGT 1946
Db      |||
QY 700 GCTGTTTATATATGGTGGTCCCTCAGGTGAGTGGTGAATATCGGTTTAAAGAGTCAA 759
Db      |||
QY 1947 GCTGTTTATATATGGTGGTCCCTCAGGTGAGTGGTGAATATCGGTTTAAAGAGTCAA 2006
Db      |||
QY 760 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAG 819
Db      |||
QY 2007 GTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAAACAG 2066
Db      |||
QY 820 GGGATCCTGTGACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 879
Db      |||
QY 2067 GGGATCCTGTGACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAAT 2126
Db      |||
QY 880 AGAAATTCAGATCAGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGA 939
Db      |||
QY 2127 AGAAATTCAGATCAGTGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGA 2186
Db      |||
QY 940 CTTAGATCGTGGGATCCAGGCTGGTCTTATCGAGGATACCTCTCCCTGATGGCAT 999
Db      |||
QY 2187 CTTAGATCGTGGGATCCAGGCTGGTCTTATCGAGGATACCTCTCCCTGATGGCAT 2246
Db      |||
QY 1000 AATGACAGATCAGATCTTTCAGGCTTCTATTCCTGGGCCCCAGTCACTCTGTGGAT 1059
Db      |||
QY 2247 AATGACAGATCAGATCTTTCAGGCTTCTATTCCTGGGCCCCAGTCACTCTGTGGAT 2306
Db      |||
QY 1060 CTTCTATGATACAGGATACACGGA 1083
Db      |||
QY 2307 CTTCTATGATACAGGATACACGGA 2330
Db      |||

RESULT 2
US-11-151-601-19
; Sequence 19, Application US/11151601
```

```
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MP100-054P1RCPLOMIDVIM
; CURRENT APPLICATION NUMBER: US/11/151.601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(2874)
; US-11-151-601-19

Query Match      73.1%; Score 791.2; DB 17; Length 3143;
Best Local Similarity 99.0%; Pred. No. 6.1e-208;
Matches 796; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 280 TGAAGGAGGATCATAGTTCATGATCAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 339
Db      |||
QY 1755 TGGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGA 1814
Db      |||
QY 340 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCGAGAGGTCACAAAG 399
Db      |||
QY 1815 CTCCTCTTTAGAGCATCACCTGTACGTAGTACGTAAATCTCGAGAGGTCACAAAG 1874
Db      |||
QY 400 GCTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGTCAGTACGACTGCTGACTTCTTTAT 459
Db      |||
QY 1875 GCTGACTGACCGTGGCTACTCACATCTTCTGTCATCAGTCAGTACGACTGCTGACTTCTTTAT 1934
Db      |||
QY 460 AAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 519
Db      |||
QY 1935 AAGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCC 1994
Db      |||
QY 520 TGAAGATGACCCAACTTGCAGAAATTTTGGGCCACCATTTTGGATTCAGCAGG 579
Db      |||
QY 1995 TGAAGATGACCCAACTTGCAGAAATTTTGGGCCACCATTTTGGATTCAGCAGG 2054
Db      |||
QY 580 TCCTCTTCTGACTACTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTTAC 639
Db      |||
QY 2055 TCCTCTTCTGACTACTCTCCAGAAATTTTCTCTTTTGAAGAGTACTACTGGAATTTAC 2114
Db      |||
```

QY 640 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGT 699  
DB 2115 ATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGT 2174  
QY 700 GCTGTTTCATATATGTTGGTCTCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGAGTCAA 759  
DB 2175 GCTGTTTCATATATGTTGGTCTCAGGTGCAAGTTGGTGAATAATCGGTTTAAAGAGTCAA 2234  
QY 760 GTATTTTCGCTTGAATACCTTAGCCTCTTAGGTTATGCTTGTAGTGTAGTATAGACAAACAG 819  
DB 2235 GTATTTTCGCTTGAATACCTTAGCCTCTTAGGTTATGCTTGTAGTGTAGTATAGACAAACAG 2294  
QY 820 GGGATCCGTGTCACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAT 879  
DB 2295 GGGATCCGTGTCACCGAGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAT 2354  
QY 880 AGAAATTCAGCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 939  
DB 2355 AGAAATTCAGCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCATTGA 2414  
QY 940 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 999  
DB 2415 CTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCAAT 2474  
QY 1000 AATGACAGGTCAGATATCTTTCAGGGTTGCTATGTCGGGGCCCAAGTCACTCTGTGGAT 1059  
DB 2475 AATGACAGGTCAGATATCTTTCAGGGTTGCTATGTCGGGGCCCAAGTCACTCTGTGGAT 2534  
QY 1060 CTTCTATGATACAGGATACACGGA 1083  
DB 2535 CTTCTATGATACAGGATACACGGA 2558

RESULT 3

US-09-925-065A-669313/c  
; Sequence 669313, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wahg, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 669313  
; LENGTH: 1346  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-669313

Query Match 10.9%; Score 118; DB 7; Length 1346;  
Best Local Similarity 100.0%; Pred. No. 5.4e-22;  
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 576 CAGGTCTCTTCTCCGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAT 635  
DB 118 CAGGTCTCTTCTCCGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAT 59  
QY 636 TTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCC 693

DB 58 TTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCC 1  
RESULT 4  
US-11-079-463-1186  
; Sequence 1186, Application US/11079463  
; Publication No. US20060073161A1  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRA  
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/11/079,463  
; CURRENT FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/128,705  
; PRIOR FILING DATE: 1999-04-09  
; PRIOR APPLICATION NUMBER: US 09/540,209  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 10444  
; SEQ ID NO 1186  
; LENGTH: 2238  
; TYPE: DNA  
; ORGANISM: B.fragilis  
US-11-079-463-1186

Query Match 6.3%; Score 68; DB 18; Length 2238;  
Best Local Similarity 49.6%; Pred. No. 4.3e-08;  
Matches 238; Conservative 0; Mismatches 230; Indels 12; Gaps 2;

QY 604 AGAAATTTCTCTTTGAAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAGCC 663  
DB 1491 AGAATTTCTTACGTTTCAAAACACAGGAAGAGTGCATCTGAACGGCTGGATGATGAAC 1550  
QY 664 TCATGATCTACAGCCTGGAAGAAATATCTCTGCTGCTGCTTCATATATGGTGGTCTCA 723  
DB 1551 GGTCAATTTTCGATCTCTGCCAAACGTTATCCGGTACTGATGTTCCAGTATAGCGGTC 1610  
QY 724 GGTGAGTTGTTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCTAGC 783  
DB 1611 TTCGCAACAGGTTCTGGACAAATGGGGAATCAGTTGGGAAACCTCATG-----GC 1661  
QY 784 CTCTCTAGTTATGTTGGTTGTAGTGATACACACAGGGGATCTGTCAACCGAGGGCTTAA 843  
DB 1662 GAGCCTCGGTTACGTTGGTAGCTTGTGTAGATGTCGGGCACAGGTGGCGGTGGCAGTGA 1721  
QY 844 ATTTGAAGCGCCTTTAAATATAAAATGAGTCAAAATAGAAATTTGACGATCAGGTGGAAG 903  
DB 1722 ATTCAGAAATGCACCTACCTGAACTGGTGTAAAGAGCTTAAAGACCCAGGTGGAAGC 1781  
QY 904 ACTCAATATCTAGCTTCTCGATATGATTTTCATTTGATCTAGATCTGTTGGGATCCACGG 963  
DB 1782 TGCCAAATATCTGGGTGGACTTGCCTTATGTG---GACAAAGGACGTTATTTGGTATCTGG 1838  
QY 964 CTGGTCTTATGGAGGATACCTCTCCCTGATGGGATTAATGACAGAGTTCAGATCTTCAG 1023  
DB 1839 ATGGAGTTTCGGCGGATATATGACCAATCATGATGATGAGCGAAGGTACACCCGTGTTAA 1898  
QY 1024 GGTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083  
DB 1899 AGCGGAGTTCTGTGGCCCGACCTACAGACTGGAATATTAGCATACAGTATATACCGA 1958

RESULT 5

US-11-208-288-3  
; Sequence 3, Application US/11208288  
; Publication No. US20060051366A1  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, Chiwen  
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF  
; FILE REFERENCE: ANGIOGENESIS AND INFLAMMATION  
; CURRENT APPLICATION NUMBER: US/11/208,288  
; CURRENT FILING DATE: 2005-08-18



```

; SEQ ID NO 34
; LENGTH: 3407
; TYPE: DNA
; ORGANISM: Homo
US-10-501-035-34

```

	Query Match	5.5%	Score 59.6;	DB 11;	Length 3407;
	Best Local Similarity	53.6%;	Pred. No. 1.1e-05;		
	Matches 147;	Conservative 0;	Mismatches 124;	Indels 3;	Gaps 1;
Qy	810	TAGACAACAGGGGATCTGTCCACGAGGGCTTAAATTGGAAGCGCCCTTTAAATATATAAA	869		
Db	1808	TTGATGGCAGAGGAAGTGGTTACCAAGSAGATAAGATCATGTGATGCAATCAACAGAGAC	1867		
Qy	870	TGGGTCAAAATAGAAAATTGACATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATG	929		
Db	1868	TGGGACATTTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAAATGG	1924		
Qy	930	ATTTTCATTGACTTAGTGTGTGGGCATCCACGGCTGGTCTCTATGAGGATACCTCTCCC	989		
Db	1925	GATTTGTGGACAACAAACGAATTGCAATTTTGGGCTGGTCATATGAGGGTACGTAACT	1984		
Qy	990	TGATGGCATTAATGCAAGGTCAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCCAGTCA	1049		
Db	1985	CAATGTGCTCGGATCGGGAAGTGGCGTGTTCAAAGTGTGAATAGCCGTGGCGCCCTGTAT	2044		
Qy	1050	CTCTGTGGATCTTCTATGATACAGGATACCGGA	1083		
Db	2045	CCCGGTGGGAGTACTATGACTCAGTGTACAGA	2078		

```

RESULT 9
US-11-136-527-2130
/ Sequence 2130, Application US/11136527
/ Publication No. US20050287570A1
/ GENERAL INFORMATION:
/ APPLICANT: Wyeth
/ APPLICANT: Mounts, William M
/ TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
/ FILE REFERENCE: 031896-041000 (AM101086)
/ CURRENT APPLICATION NUMBER: US/11/136,527
/ CURRENT FILING DATE: 2005-05-25
/ PRIOR APPLICATION NUMBER: US 60/574,294
/ PRIOR FILING DATE: 2005-05-26
/ NUMBER OF SEQ ID NOS: 362830
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 2130
/ LENGTH: 4852
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
US-11-136-527-2130

```

	Query Match	5.4%;	Score 58.4;	DB 17;	Length 4852;
	Best Local Similarity	51.4%;	Pred. No. 2.6e-05;		
	Matches 147;	Conservative	5;	Mismatches 131;	Indels 3; Gaps 1;
Qy	798	TGGTGTAGTGATAGACACACAGGGGATCCTGTCCAGGAGGCCTAAATTTGAAGGCGCCT	857		
Db	1812	TRCTAGCTAGCTTTTGATGGCAGKAAGTGGTTTACCAGGAGATTAAGATCATGCAATGCCAA	1871		
Qy	858	TTAATAATAAAATGGGTCAAATATAGAAAAATTGACGATCAGGTGGAAGGACTCCAATATATCTAG	917		
Db	1872	TCAACAAAAGACTTGGNACACTGGAAAGTTGAAGATCAAAATTTGAAGCAGCC---AGGCAAT	1928		
Qy	918	CTTCTCGATATGATTTTCATTGACTTAGATTCGTGTGGGCATCCA CGGCTGGTCTTATCGAG	977		
Db	1929	TTTTAAAAATGGGATTTGTGGACAGCAACRAGTTGCAATTTGGGCGCTGGTCATATCGAG	1988		
Qy	978	GATACCTCTCCCTGATGGCATTAAATGCAGAGCTCAGATATCTTCAGGGTTGCTATTGCTG	1037		
Db	1989	GGTAGCTGAACCTCAATGGTCTCTGGGATCGGAAGTGGCGTGTTCGAARTGTGGAATACCG	2048		
Qy	1038	GGGCCCCAGTCACTCTGTGGATCTTCTTATGATACAGGATACACGA	1083		

db  
2049 TGGCGCCCGTGCACGGTGGGAGTACTATGACTCAGTATACACAGA 2094

```

RESULT 10
US-11-079-463-2536
; Sequence 2536, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRAGILIS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 2536
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: B.fragilis
US-11-079-463-2536

```

Query Match	5.1%;	Score 54.8;	DB 18;	Length 1884;
Best Local Similarity	47.6%;	Pred. No. 0.00018;		
Matches 195;	Conservative 0;	Mismatches 212;	Indels 3;	Gaps 1;
QY	659	AAGCCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTGTGTTCTATATATATGTTGGT	718	
Db	1186	AAACCGCGCGATTTTCGATCCGAACAAGAAATACCCGGCTATCGTATACGTATATACGGCGGT	1245	
QY	719	CCTCAGGTGCGAGTTGGTGAATAATATCGGTTTAAAGGAGTCAAGTATTTCCCGTTGATATACC	778	
Db	1246	CCTCAGCGCACAACTGGTCAACAACGGCTGGCGAAGCGGTGCACGAGGCTGGGATATCTAT	1305	
QY	779	CTAGCGCTCTTAGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCCTGTCAACCGAGGG	838	
Db	1306	ATGGCCAAACAAGGTTATCATCATGTTTACCGTAGACGACGCTGGAAGCAGCAATTCGGGGA	1365	
QY	839	CTTAAATTGGAAGCGCCCTTTAAATAATAAAATGGGTCAATAGAGAAATTGACGTACAGTG	898	
Db	1366	CTCGATTTTGAGATGTTTACTTTCCGCGAGTTGGGAATCGAAGAAGGAAGACACAGAGTG	1425	
QY	899	GAAGGACTCCAAATCTAGCTTCTCGATATGATATTTCATTGACTTTAGATCGTGTGGGCATC	958	
Db	1426	AAAGGAACCGAATTCCTGAAAGCCTCCGCTACGTGATGGGA---ACCGTATCCGAGTA	1482	
QY	959	CACGGCTGGTCTCATGGAAGGATACCTCTCTCCGTGATGGCAATTAATGACGAGGTCAAGATATC	1018	
Db	1483	CACGGCTGGAGCTTTGGGGGTCACATGAACAACGTCGCCCTTCTCTTCGTATATCCCGAGATA	1542	
QY	1019	TTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGA	1068	
Db	1543	TTTAAAGTCGGCGTGGCGGGCGGTCCCGTCATTGACTGGGGTATTACGA	1592	

RESULT 11  
US-11-208-288-5  
; Sequence 5, Application US/11208288  
; Publication No. US20060051366A1  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, Chiwen  
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF  
; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION  
; FILE REFERENCE: 39533-0001  
; CURRENT APPLICATION NUMBER: US/11/208,288  
; CURRENT FILING DATE: 2005-08-18  
; PRIOR APPLICATION NUMBER: US 60/605,013  
; PRIOR FILING DATE: 2004-08-26  
; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2283  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-11-208-288-5

Query Match  
Best Local Similarity 51.8%; Score 54.8; DB 18; Length 2283;  
Matches 144; Conservative 0; Mismatches 127; Indels 3; Gaps 1;  
QY 810 TAGACACAGGGAGTCTGTGACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAA 869  
DB 1715 TTGACGCGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGAT 1774  
QY 870 TGGGTCAAAATAGAAATGACGATGAGGAGTCCATATCTAGCTTCTCGATG 929  
DB 1775 TGGGAACACTGGAAGTTGAAGATCAAAATGAGCAGCC---AGCAATTTGTAATG 1831  
QY 930 ATTTCATTGACTTAGATCTGTGGGCATCCAGCGTGTCTCTATGGAGGATACCTCTCC 989  
DB 1832 GATTGTGGATCAAGCAGTTGCNAATTTGGGGCTGTCTATGAGGGGTATGTAACCT 1891  
QY 990 TGATGCGATTAATGACAGGTGAGATATCTTCAGGTTGCTATTGCTGGGCCCCAGTCA 1049  
DB 1892 CAATGCTCTGGATCGGAAGTGGCGTTTCAAGTGGGAATAGCTGTGGCACCCTGTGT 1951  
QY 1050 CTCTGGGATCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1083  
DB 1952 CACGGTGGAGTACTATGACTCAGTGTGTACACAGA 1985

## RESULT 12

US-11-226-869-428  
; Sequence 428, Application US/11226869  
; Publication No. US20060069054A1  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.470C14  
; CURRENT APPLICATION NUMBER: US/11/226.869  
; NUMBER OF SEQ ID NOS: 627  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 428  
; LENGTH: 538  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-226-869-428

Query Match  
Best Local Similarity 4.8%; Score 52.4; DB 18; Length 535;  
Matches 150; Conservative 0; Mismatches 141; Indels 3; Gaps 1;  
QY 790 AGGTATGTGTTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849  
DB 137 AGGATGTCTATCCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 196  
QY 850 AGGCGCTTTAAATATAAAATGAGTCAAAATGAGTCAAAATGAGTCAAAATGAGTCAAA 909  
DB 197 CTATGCAAGTATCGAAGCTGGGTGTTTATGAAGTTGAAGACCAAGATTACAGCTGTG 256  
QY 910 ATATCTAGCTTCTCGATATGATTTCAATGATGATGATGATGATGATGATGATGATGAT 969  
DB 257 AAA---ATTCAAGAAATGGTTTCAATGATGATGATGATGATGATGATGATGATGATGAT 313  
QY 970 CTATGAGGATACCTCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029  
DB 314 CTATGAGGATACCTCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 373

QY 1030 TATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083  
DB 374 TATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTACGCGTCTGTCTACACAGA 427

## RESULT 13

US-10-505-928-476  
; Sequence 476, Application US/10505928  
; Publication No. US2006008532A1  
; GENERAL INFORMATION:  
; APPLICANT: Ludwig Institute for Cancer Research et al.  
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
; FILE REFERENCE: 28967/39178  
; CURRENT APPLICATION NUMBER: US/10/505.928  
; PRIOR FILING DATE: 2004-08-27  
; PRIOR APPLICATION NUMBER: US 60/363.019  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: PatentIn 3.2  
; SEQ ID NO 476  
; LENGTH: 2788  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-505-928-476

Query Match  
Best Local Similarity 4.8%; Score 52.4; DB 9; Length 2788;  
Matches 150; Conservative 0; Mismatches 141; Indels 3; Gaps 1;  
QY 790 AGGTATGTGTTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849  
DB 1903 AGGATGTCTATCCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1962  
QY 850 AGGCGCTTTAAATATAAAATGAGTCAAAATGAGTCAAAATGAGTCAAAATGAGTCAAA 909  
DB 1963 CTATGCAAGTATCGAAGCTGGGTGTTTATGAAGTTGAAGACCAAGATTACAGCTGT 2022  
QY 910 ATATCTAGCTTCTCGATATGATTTCAATGATGATGATGATGATGATGATGATGATGAT 969  
DB 2023 AAA---ATTCAAGAAATGGGTTCATGATGATGATGATGATGATGATGATGATGATGAT 2079  
QY 970 CTATGAGGATACCTCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029  
DB 2080 CTATGAGGATACCTCTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2139  
QY 1030 TATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 1083  
DB 2140 TATAGCAGTGGCTCCAGTCTCCAGTGGGAATATTACGCGTCTGTCTACACAGA 2193

## RESULT 14

US-11-186-284-54  
; Sequence 54, Application US/11186284  
; Publication No. US2005026693A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burgart, Lawrence J.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; FILE REFERENCE: MP001-039P2RNM  
; CURRENT APPLICATION NUMBER: US/11/186.284  
; PRIOR FILING DATE: 2005-07-21  
; PRIOR APPLICATION NUMBER: US/10/301.822  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339.971  
; PRIOR FILING DATE: 2001-12-10

```

, PRIOR APPLICATION NUMBER: US 60/361,978
, PRIOR FILING DATE: 2002-03-05
, PRIOR APPLICATION NUMBER: US 60/381,988
, PRIOR FILING DATE: 2002-05-20
, NUMBER OF SEQ ID NOS: 228
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 54
, LENGTH: 2814
, TYPE: DNA
, ORGANISM: Homo Sapiens
, FEATURE:
, NAME/KEY: CDS
, LOCATION: (209)...(2491)
US-11-186-284-54

```

Query Match	4.8%;	Score 52.4;	DB 17;	Length 2814;
Best Local Similarity	51.0%;	Pred. No. 0.00097;		
Matches 150;	Conservative 0;	Mismatches 141;	Indels 3;	Gaps 1;
790	AGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTACCCAGGAGGCTTAAATTTGA	849		
1903	AGGGATGGTCAATTGGCTTGGTGGATGTCAGGAAACAGCTTTCAGAGGTGACAACTCCT	1962		
850	AGGGCCCTTTAAATAATAAATAGGCTCAAAATAGAAATTGACGATCAGGTGGAGGACTCCA	909		
1963	CTATGCAGGTGATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGATTACAGCTGTGAG	2022		
910	ATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATTCGTGTGGGATCCAGCGCTGTC	969		
2023	AAA---ATTCTAAGAAATGGGTTTTCAATGATGAAAAAAGAAATAGCCATATGGGGCTGGTC	2079		
970	CTATGGAGGATACCTTCCCTGTAGGCCATTAAATGCAGAGTCAGATATCTTCAGGGTTGC	1029		
2080	CTATGGAGGATACGTTTTATCACTGGGCCCTTGCATCTGGAACTGGTCTTTTCAATGTGG	2139		
1030	TATTGCTGGGGCCCAAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGA	1083		
2140	TATAGCAGGTGGCTCCAGTCTCCAGCTGGGAATATTACCGCTGTCTACACAGA	2193		

Search completed: May 4, 2006, 09:26:09  
Job time : 1481.29 secs

RESULT 15  
US-11-245-147-168  
; Sequence 168, Application US/11245147  
; Publication No. US20060030541A1  
; GENERAL INFORMATION:  
; APPLICANT: GARCIA, TERESA  
; APPLICANT: ROMAN ROMAN, SERGIO  
; APPLICANT: BARON, ROLAND  
; APPLICANT: CALL, KATHERINE  
; APPLICANT: THEILHABER, JOACHIM  
; APPLICANT: CONNOLLY, TIMOTHY  
; APPLICANT: JACKSON, AMANDA  
; APPLICANT: BUSHNELL, STEVEN  
; APPLICANT: RAWADI, GEORGES  
; TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE  
; FILE REFERENCE: 37991-0023  
; CURRENT APPLICATION NUMBER: US/11/245,147  
; CURRENT FILING DATE: 2005-10-07  
; PRIOR APPLICATION NUMBER: PCT/IB02/02211  
; PRIOR FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: 60/281,400  
; PRIOR FILING DATE: 2001-04-05  
; NUMBER OF SEQ ID NOS: 246  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 168  
; LENGTH: 2814  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Homo sapiens fibroblast activation protein, alpha  
; OTHER INFORMATION: (FAP), mRNA  
US-11-245-147-168



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocolloration Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: May 2, 2006, 00:41:48 ; Search time 32.8222 Seconds  
(without alignments)  
2894.257 Million cell updates/sec

Title: US-10-825-632-8  
Perfect score: 1938  
Sequence: 1 ggaagaagatgccagatcag.....tatgatacaggatacacgga 1083

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/abss/ABSSWEB.spool/US10825632/runat\_01052006\_105944\_3214/app.query.fasta.1  
-DB=A.Geneseq -QFNT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p  
-USER=US10825632@cgn 1.1.605 @runat\_01052006\_105944\_3214 -NCPU=6 -ICPU=3  
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.Geneseq 21:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*
- 9: Geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1933	99.7	360	4 AAB47190	AAB47190 Human DPP
2	1836.5	94.8	824	6 ABU92030	ABU92030 Human pro
3	1836.5	94.8	882	4 AAB47187	AAB47187 Human DPP
4	1836.5	94.8	882	5 ABG61591	ABG61591 Human DPP
5	1836.5	94.8	882	5 AAG78415	AAG78415 Amino aci
6	1836.5	94.8	882	5 AAE24170	AAE24170 Human dip
7	1836.5	94.8	882	5 AAU74749	AAU74749 Human pro
8	1836.5	94.8	882	5 ADI17086	ADI17086 Human NOV
9	1836.5	94.8	882	6 ABU07720	ABU07720 Human ser

10	1756.5	90.6	883	5 ADI17085	ADI17085 Murine NO
11	1534	79.2	831	6 ABU92026	ABU92026 Human pro
12	1533.5	79.1	493	7 ADE78977	ADE78977 Human pro
13	1528	78.8	632	4 AAB93565	AAB93565 Human pro
14	1525.5	78.7	587	5 ADR41398	ADR41398 Human CD-
15	1495	77.1	580	5 AAE14337	AAE14337 Human pro
16	1466.5	75.7	746	8 ADI16334	ADI16334 Human pro
17	1254.5	64.7	724	5 ABB97362	ABB97362 Novel hum
18	1254.5	64.7	782	5 ABB97361	ABB97361 Novel hum
19	1226	63.3	690	5 ABG61594	ABG61594 Human DPR
20	1220.5	63.0	689	5 ADI17084	ADI17084 Human NOV
21	1220.5	63.0	830	5 AAE24171	AAE24171 Human dip
22	1220.5	63.0	863	5 ABG61592	ABG61592 Human DPP
23	1220.5	63.0	863	5 ADI17083	ADI17083 Human NOV
24	1220.5	63.0	892	5 ABG61602	ABG61602 Human DPR
25	1220.5	63.0	892	5 ABG61604	ABG61604 Human DPR
26	1220.5	63.0	892	5 ABB98134	ABB98134 Human PMM
27	1220.5	63.0	892	8 ADS10951	ADS10951 Human the
28	1220.5	63.0	969	5 AAE24168	AAE24168 Human dip
29	1219	62.9	516	6 ABU92029	ABU92029 Human pro
30	1213.5	62.6	755	8 ADQ67811	ADQ67811 Novel hum
31	1212.5	62.6	863	5 ADI16690	ADI16690 Human NOV
32	1212.5	62.6	863	5 ADI16688	ADI16688 Human NOV
33	1212.5	62.6	863	8 ADN42344	ADN42344 Human nov
34	1209	62.4	661	5 ABG61596	ABG61596 Human DPR
35	1204.5	62.2	847	5 AAE23875	AAE23875 Murine di
36	1204.5	62.2	869	5 AAE24169	AAE24169 Alternati
37	1201	62.0	879	5 ABG61607	ABG61607 Human DPR
38	1201	62.0	879	5 ABG61608	ABG61608 Human DPR
39	1201	62.0	879	8 ADS10953	ADS10953 Human the
40	1199.5	61.9	658	5 ABG61600	ABG61600 Human DPR
41	1196.5	61.7	863	8 ADN42342	ADN42342 Human nov
42	1183.5	61.1	683	4 AAM40510	AAM40510 Human pol
43	1183.5	61.1	737	4 AAM38724	AAM38724 Human pol
44	1183.5	61.1	873	8 ADS10479	ADS10479 Human the
45	1174.5	60.6	854	8 ADS10952	ADS10952 Human the

#### ALIGNMENTS

RESULT 1  
AAB47190

ID : AAB47190 standard; protein; 360 AA.

XX  
AC AAB47190;  
XX  
DT 29-JUN-2001 (first entry)  
XX  
DE Human DPP8 244Glu-341Ile+515Val-776Thr.  
XX  
KW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;  
KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;  
KW growth hormone deficiency; Glucose level; mucosal regeneration;  
KW non-insulin dependent diabetes mellitus; glucose intolerance;  
KW immunosuppression.

XX Homo sapiens.

XX WO200119866-A1.

XX PD 22-MAR-2001.

XX PF 11-SEP-2000; 2000WO-AU001085.

XX PR 10-SEP-1999; 99AU-00002762.

XX PR 18-FEB-2000; 2000AU-00005709.

XX PA (UNSY ) UNIV SYDNEY.

XX PI Abbott CA, Gorell MD;

XX XX WPI; 2001-281520/29.

DR N-PSDB; AAC85697.



PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;  
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
XX WPI; 2003-430274/40.  
DR N-PSDB; ACA92425.  
XX  
PT New human protein modification and maintenance molecules (PMMM), useful  
PT for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or  
PT infections.

XX Claim 1; Page 249-251; 31pp; English.

XX The present invention relates to the isolation of human protein  
CC modification and maintenance molecules (PMMM), and the polynucleotide  
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM  
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful  
CC for diagnosing a condition or disease associated with the expression of  
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and  
CC generating an expression profile of a sample containing the  
CC polynucleotides. The diseases or conditions associated with decreased  
CC expression or overexpression of PMMM are cell proliferation disorders  
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,  
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
CC syndrome), gastrointestinal or epithelial disorders, and infections. The  
CC PMMM polypeptides or their fragments are useful in screening compounds  
CC for effectiveness as agonists or antagonists of the polypeptides, or in  
CC altering the expression of the target polynucleotide and compounds that  
CC specifically bind to, or modulate the activity of the polypeptide.  
CC ABU92021-ABU92060 represent the human PMMM polypeptides of the invention

XX SQ Sequence 824 AA;

Alignment Scores:  
Pred. No.: 1, 7e-204 Length: 824  
Score: 1836.50 Matches: 360  
Percent Similarity: 67.5% Conservative: 0  
Best Local Similarity: 67.5% Mismatches: 0  
Query Match: 94.8% Indels: 173  
DB: 6 Gaps: 1

US-10-825-632-8 (1-1083) x ABU92030 (1-824)

QY 2 GAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGA 61  
DB 186 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 205  
QY 62 TATTCTGGCTATTGGTGGTCCAAAGCTGAAACAACTCCCGAGTGGTGAATTTCTT 121  
DB 206 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 225  
QY 122 AGAATTCATATGAGAAATATGATGATCTGAGGTGGAAATTTATCATGTTACATCCCT 181  
DB 226 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 245  
QY 182 ATGTGCAACAGAGGGGAGATTCATTCCGTTATCTTAAACAGGTACAGCAATCCT 241  
DB 246 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 265  
QY 242 AAAGTCACTTTTAAAGATGTGAGAAATAATGATTGATGCTGAGGAGGATCATTA 295  
DB 266 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 285  
QY 295 ----- 295  
DB 286 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValcIuTyrIleAla 305  
QY 295 ----- 295  
DB 306 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 325  
QY 295 ----- 295

DB 326 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 345  
QY 295 ----- 295  
DB 346 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 365  
QY 295 ----- 295  
DB 366 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 385  
QY 295 ----- 295  
DB 386 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 405  
QY 295 ----- 295  
DB 406 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 425  
QY 295 ----- 295  
DB 426 AlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSerGlyGluTrp 445  
QY 296 ----- GTTGATGAAGTCAGAGGCTGTATAT 322  
DB 446 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 465  
QY 323 TTTGAGGCGACCAAGATCCCTTTAGACATCACCCTGACTAGTACAGTACGTAAT 382  
DB 466 PheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSerTyrValAsn 485  
QY 383 CCTGAGAGGTGACAGGCTGACTGACCTGGCTACTACATCTTGCTGCATCAGTCAG 442  
DB 486 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysLysIleSerGln 505  
QY 443 CACTGTGACTTCTTTATAGTATAGTAAGTAACTCCAGAGAAATTCACACTGTGTCCCTT 502  
DB 506 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 525  
QY 503 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTGTCACAAACAAAGAAATTTGGGCCACC 562  
DB 526 TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 545  
QY 563 ATTTTGGATTCAGCAGGCTCTCTCTGACTATACCTCCAGAAATTTCTCTTTTCAA 622  
DB 546 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 565  
QY 623 AGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 682  
DB 566 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 585  
QY 683 AAGAAATATCCTACTGTGCTGTTTCAATATATGGTGGTCTCAGGTGCAGTGTGTGAATAT 742  
DB 586 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 605  
QY 743 CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTACCTCTCTAGTTATGTGTT 802  
DB 606 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 625  
QY 803 GTAGTGATAGACAACAGGGGATCTCTCACCAGGGCTTAAATTTGAGGCCCTTTAAA 862  
DB 626 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluIleAlaPheLys 645  
QY 863 TATAAATGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCT 922  
DB 646 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 665  
QY 923 CGATATGATTTTCAATGATCTAGTCTGTGGGCATCCAGGCTGTGCTTATGAGGATAC 982  
DB 666 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrTrpSerTyrGlyGlyTyr 685  
QY 983 CTCCTCCCTGATGGCATTAAATGACAGTCAATCTTCAGGGTTCCTATTCCTGGGCC 1042

Db 686 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 705  
 QY 1043 CCAGTCACTCTGGGATCTCTATGATACAGATACAG 1081  
 Db 706 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 718

## RESULT 3

AA047187  
 ID AAB47187 standard; protein; 882 AA.

AC AAB47187;

DI 29-JUN-2001 (first entry)

DE Human DPP8.

KW Human; dipeptidyl aminopeptidase; DPP8; prolly oligopeptidase;  
 KW dipeptidyl peptidase; DPP8; T cell; cleavage; diarrhoea;  
 KW growth hormone deficiency; glucose level; mucosal regeneration;  
 KW non-insulin dependent diabetes mellitus; glucose intolerance;  
 KW immunosuppression.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"  
 FT Active-site 817  
 FT /note= "Forms part of Ser-Asp-His catalytic triad"  
 FT Active-site 849  
 FT /note= "Forms part of Ser-Asp-His catalytic triad"

PN W0200119866-A1.

PD 22-MAR-2001.

PF 11-SEP-2000; 2000WO-AU001085.

PR 10-SEP-1999; 99AU-00002762.

PR 18-FEB-2000; 2000AU-00005709.

PA (UNSY ) UNIV SYDNEY.

PI Abbott CA, Gorell MD;

DR WPI; 2001-281520/29.

DR N-PSDB; AAC85694.

FT New human dipeptidyl aminopeptidase (DPP8) useful for cleaving  
 FT substrates, identifying inhibitors of DPP8 catalytic activity which have  
 FT therapeutic uses, and for detecting activated T cells.

PS Claim 1; Fig 2; 78pp; English.

CC This sequence represents human dipeptidyl aminopeptidase (DPP8). DPP8 has  
 CC substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.  
 CC Therefore, it is a prolly oligopeptidase and a dipeptidyl peptidase.  
 CC Because it is capable of hydrolysing the peptide bond C-terminal pro  
 CC in each of these compounds, DPP8 is homologous with human DPPIV. DPP8 is  
 CC useful for cleaving a substrate, and for detecting an activated T cell  
 CC which involves measuring the level of DPP8 gene expression in a T cell.  
 CC The level of DPP8 expression is detected by detecting the amount of DPP8  
 CC RNA in the cell. It is also useful for identifying a molecule capable of  
 CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as  
 CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,  
 CC growth hormone deficiency, lowering glucose levels in non-insulin  
 CC dependent diabetes mellitus and other disorders involving glucose  
 CC intolerance, enhancing mucosal regeneration and as immunosuppressants

XX Sequence 882 AA;

SQ Alignment Scores:

Pred. No.: 1.75e-204 Length: 882

Score: 1836.50 Matches: 360  
 Percent Similarity: 67.5% Conservative: 0  
 Best Local Similarity: 67.5% Mismatches: 0  
 Query Match: 94.8% Indels: 173  
 DB: 4 Gaps: 1  
 US-10-825-632-8 (1-1083) x AAB47187 (1-882)  
 QY 2 GAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGAAATTTGATAGA 61  
 Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 263  
 QY 62 TATTCTGGCTATTGGTGTCTCCAAAAGCTGAAACAACTCCAGAGTGGTGGTAAAAATCTT 121  
 Db 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283  
 QY 122 AGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAAATATTTCATGTTACATCCCT 181  
 Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303  
 QY 182 ATGTTGAAACAAAGGAGGCGAGATTCATTCCCTTATCTTAAACACAGGTACAGCAATCCT 241  
 Db 304 MetLeuGluThrArgArgAlaAspSerPheA-TyrProLysThrGlyThrAlaAsnPro 323  
 QY 242 AAAGTCACTTTAAGATGTCAGAAATATGATGATGCTGAAGGAAGGATCATATA----- 295  
 Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343  
 QY 295 ----- 295  
 Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363  
 QY 295 ----- 295  
 Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383  
 QY 295 ----- 295  
 Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403  
 QY 295 ----- 295  
 Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423  
 QY 295 ----- 295  
 Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443  
 QY 295 ----- 295  
 Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463  
 QY 295 ----- 295  
 Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483  
 QY 295 ----- 295  
 Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerSerGlyGluTrp 503  
 QY 296 -----GTTGATGAAGTCAGAGGCTCGTATAT 322  
 Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523  
 QY 323 TTTGAAGGCACCAAGACTCCCTTTAGACATACCTGTACGTAGTCAGTACGTAAAT 382  
 Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543  
 QY 383 CTGTGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCTTGTCTGATCAGTCAG 442  
 Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563  
 QY 443 CACTGTGACTTCTTTATAGTAAAGTATAGTAACACAGAGAAGATCCACACTGTGTCTCCTT 502

```
Db 564 HisCysaspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 503 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTCGAAACAAAGAAATTTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhePheAlaThr 603
Qy 563 ATTTTGGATTACAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
Qy 623 AGTACTACTGTGATTACATTGATGGATGCTCTCAAGCCCTCATGATCTACAGCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 683 AAGAAATATCTCTCTGCTGTTTCATATATGTGGTCTCTCAGGTGCAGTTGGTGAATAAT 742
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 743 CGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCCTCTCTAGGTTATGTGGTT 802
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 803 GTACTGATACACACAGGGATCTCTCACCGAGGCTTAAATTTGAAGCGCTTTAAA 862
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyValAlaPheLys 703
Qy 863 TATAAATGGTCAAAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 923 CGATATGATTTCAATGATTTAGATCTGTGGGATCCACGGCTGCTCTATGAGGATAC 982
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
Qy 983 CTCTCCCTGATGGCATTATGACAGGTGATATCTTCAGGTTGCTATTGCTGGGGCC 1042
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 1081
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

RESULT 4
ABG61591
ID ABG61591 standard; protein; 882 AA.
XX
AC ABG61591;
XX
DT 12-AUG-2002 (first entry)
XX
DE Human DPPIV related serine protease DPPP-1.
XX
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX
OS Homo sapiens.
XX
PN WO200231134-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US031874.
XX
PR 12-OCT-2000; 2000US-0240117P.
XX
PA (FERR ) FERRING BV.
XX
```

```
PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
DR N-PSDB; ABK83322.
XX
PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
PS Claim 17; Fig 1; 113pp; English.
XX
CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPPP proteins
XX
SQ Sequence 882 AA;

Alignment Scores:
Pred. No.: 1,75e-204 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 5 Gaps: 1

US-10-825-632-8 (1-1083) x ABG61591 (1-882)
Qy 2 GAAGAAGATGCAGATCAGCTGGAGTGGTACCTTTTCTCCAGAAGAATTTGATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
Qy 62 TATCTGGCTATTGGTGGTGTCCAAAGCTGAACCACTCCACAGTGGTGGTAAATCTT 121
Db 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283
Qy 122 AGAATTCCTATATGAAGAAATGATGATCTGAGTGAATATTTCATCTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
Qy 182 ATGTTGAAACACAGGAGGCGAGATTTCATTCCTGTTATCTTAAACAGGTACAGCAATCT 241
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
Qy 242 AAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAAGGATCATATA 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
Qy 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
Qy 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
Qy 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
Qy 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
```

QY	295	-----	295
Db	424	GluGluThrThrAspIleThrPheAenIleHisAspIlePheHisValPheProGlnSer	443
QY	295	-----	295
Db	444	HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu	463
QY	295	-----	295
Db	464	TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro	483
QY	295	-----	295
Db	484	AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp	503
QY	296	-----	322
Db	504	GluValLeuGlyArgHisGlySerAenIleGlnValAspGluValArgArgLeuValTyr	523
QY	323	TTTCGAAGGCACCAAGACTCCCTTTAGAGCATCCCTGTAGTCAGTACGTTAGTAAAT	382
Db	524	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	543
QY	383	CCTCGAGAGTCACAGGCTGACTGACCGTGGCTACTCACATCTGCTGCATCAGTCAG	442
Db	544	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln	563
QY	443	CACGTGACTCTTTATAGTAAGTATAGTAACACAGAGAAATCCACACTGTGTCTCCTT	502
Db	564	HisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSerLeu	583
QY	503	TACAAGCTATCAAGTCTGAAGATGACCCAACTTCGAAACAAAGAAATTTTGGGCCACC	562
Db	584	TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr	603
QY	563	ATTTTGGATTACAGAGTCTCTTCTCTGACTACTCTCTCCAGAAATTTTCTCTTTGAA	622
Db	604	IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu	623
QY	623	AGTACTACTGGATTTACATTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA	682
Db	624	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	643
QY	683	AAGAAATATCTACTGTGCTTTCATATATGTGTGCTCTCAGGTGCAGTCTGTCGAATAAT	742
Db	644	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn	663
QY	743	CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTT	802
Db	664	ArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGlyTyrValVal	683
QY	803	GTAAGTATACACACAGGGATCTCTGACCGAGGCTTAAATTTGAAGGCGCTTTAAA	862
Db	684	ValValIleAspAenArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	703
QY	863	TATAAATGGGTCAAATAGAAATGACATCGAGTGGAGGAGTCCCAATATCTAGCTTCT	922
Db	704	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	723
QY	923	CGATATGATTTCATCTAGATCTGTGGGCATCCACGGCTGCTCTATGGAGGATAC	982
Db	724	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr	743
QY	983	CTCTCCCTGATGGCATTATGACAGAGTTCAGATATCTTACAGGTTGCTATTGCTGGGGCC	1042
Db	744	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763
QY	1043	CCAGTCTACTCTGGGATCTCTATGATACAGATACACG	1081
Db	764	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr	776

RESULT 5	
ID	AAG78415
XX	standard; protein; 882 AA.
AC	AAG78415;
XX	
DT	12-APR-2002 (first entry)
XX	
DE	Amino acid sequence of 21953 human prolyl oligopeptidase.
XX	
KW	21953 prolyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW	cardiovascular disease; autoimmune disease; atopic allergy;
KW	neuronal disorder; vascular disorder; prostate disorder; cytostatic;
KW	antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW	diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW	Grave's disease; neuronal disorder; demyelinating disease.
XX	
OS	Homo sapiens.
XX	
PN	WO200179473-A2.
XX	
PD	25-OCT-2001.
XX	
PF	11-APR-2001; 2001WO-US040483.
XX	
PR	18-APR-2000; 2000US-0197508P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Meyers RA, Williamson M;
XX	
DR	WPI; 2002-034353/04.
DR	N-PSDB; AAH99934.
XX	
PT	New polypeptides 21953, member of human prolyl oligopeptidase family,
PT	useful as diagnostic targets and therapeutic agents for controlling
PT	cancer, lymphoma and leukemia.
XX	
PS	Claim 1; Page 102-103; 12ipp; English.
XX	
CC	This invention relates to an isolated 21953 human prolyl oligopeptidase.
CC	Which is cystostatic, antidiabetic, antiarthritic, neuroprotective,
CC	antithyroid, dermatological, antipsoriatic, antiasthmatic,
CC	ophthalmological, antiinflammatory, nootropic, antiparkinsonian,
CC	anticonvulsant, gynaecological, vasotrophic, antianginal, cardiac,
CC	antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC	gene therapy, expression or activity of 21953 protein modulator, it is
CC	useful for identifying a compound which binds to it and can be used in
CC	preventing, treating or detecting a cellular proliferative or
CC	differentiative disorder. The 21953 molecules can act as novel diagnostic
CC	targets and therapeutic agents for controlling disorders associated with
CC	the aberrant activity or degradation of peptide hormones e.g., disorders
CC	associated with cell differentiation and proliferation such as cancer,
CC	immune function, reproductive, neurological and cardiovascular function.
CC	The 21953 molecules are thus useful for treating and preventing cellular
CC	proliferative and differentiative disorders, haematopoietic neoplastic
CC	disorders, immune disorders such as autoimmune diseases, diabetes
CC	mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC	neuronal disorders, demyelinating diseases, vascular disorders and
CC	metabolism or pain disorders. This sequence represents the amino acid
CC	sequence of 21953 human prolyl oligopeptidase
XX	
SQ	Sequence 882 AA;
Alignment Scores:	
Pred. NO.:	1.75e-204
Score:	1836.50
Length:	882
Matches:	360
Percent Similarity:	67.5%
Conservative:	0
Best Local Similarity:	67.5%
Mismatches:	0
Query Match:	94.8%
Indels:	173
Gaps:	5
DB:	1
US-10-825-632-8 (1-1083) x AAG78415 (1-882)	



QY	2	GAGAAGATCCAGATCAGCTGAGTCGCTACTTGTCTCCCAAGAAATTTGATAGA	61	QY	563	ATTTTGGATTACAGAGGTCTCTTCTGACTATACCTCTCCAGAAATTTTCTTTGAA	622
Db	244	GlulGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg	263	Db	604	IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu	623
QY	62	TAATCTGGCTATTGGTGGTCTCAAAAGCTGAACAACCTCCAGTGGTGGTAAATCTT	121	QY	623	AGTACTACTGATTTACATTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTCGA	682
Db	264	TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyIleLeu	283	Db	624	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	643
QY	122	AGAATTCATATGAAGAAATGATGATCTGAGGTGGAAATTAATTCATGTTACATCCCT	181	QY	683	AAGAAATATCTACTGCTGCTTTCATATATGGTGGTCTCAGGTGAGTGTGTAATAT	742
Db	284	ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro	303	Db	644	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn	663
QY	182	ATGTGGAAACAAGAGGGAGATTCATTCCTGTTATCTCAACAGATACGCAATCCT	241	QY	743	CGTTTAAAGAGTCAAGTATTTCCGCTTCAATACCTAGCTCTCTCTAGTTATGTGTT	802
Db	304	MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro	323	Db	664	ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal	683
QY	242	AAAGTCATTTAAGATGTGAGAAATAATGATTGATGCTGAAGGAAGGATCATA	295	QY	803	GTAGTGATAGACAACAGGGATCTCTGTCACCGAGGGCTTAAATTTGAAGCGCCTTAA	862
Db	324	LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal	343	Db	684	ValValIleAspAsnAspGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	703
QY	295	-----	295	QY	863	TATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCT	922
Db	344	IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla	363	Db	704	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	723
QY	295	-----	295	QY	923	CAATATGATTTTCATTGACTTAGATCGTGGGCATCCAGCTGCTGCTATGGAGGATAC	982
Db	364	ArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln	383	Db	724	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr	743
QY	295	-----	295	QY	983	CTCTCCCTGATGGCAATTAATGCAGAGTTCAGATATCTTCAGGGTGTGCTATTGCTGGG	1042
Db	384	ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal	403	Db	744	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763
QY	295	-----	295	QY	1043	CCAGTCACTCTGTGGATCTTCTATGATACAGATACAGCAGC 1081	
Db	404	MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr	423	Db	764	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776	
QY	295	-----	295	RESULT 6			
Db	424	GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer	443	ID	AAE24170	standard; protein; 882 AA.	
QY	295	-----	295	XX	AAE24170;		
Db	444	HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu	463	XX	AC		
QY	295	-----	295	XX	DT	23-SEP-2002 (first entry)	
Db	464	TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuPro	483	XX	DE	Human dipeptidyl peptidase 8 (DPP8) protein.	
QY	295	-----	295	XX	KW	Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;	
Db	484	AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp	503	XX	KW	autoimmunity; human immuno deficiency virus; HIV infection; cystostatic;	
QY	296	-----	296	XX	KW	graft rejection; antidiabetic; antiinflammatory; immunosuppressive;	
Db	504	GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr	523	XX	OS	antiviral; enzyme.	
QY	323	TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGTAAAT	382	XX	PN	Homo sapiens.	
Db	524	PheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSerTyrValAsn	543	XX	WO200234900-A1.		
QY	383	CCTGGAGGTGACAAGGTGACTGACCGTGGTACTCACATCTTCTGTCATCAGTCAG	442	XX	02-MAY-2002.		
Db	544	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln	563	XX	29-OCT-2001; 2001WO-AU001388.		
QY	443	CATGTGACTCTTTTATAAGTATAGTATACCAAGAAATCCACACTGTGTGCTT	502	XX	27-OCT-2000; 2000AU-00001078.		
Db	564	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu	583	XX	(UNSY ) UNIV SYDNEY.		
QY	503	TACAGTATCAAGTCTCTGAGATGACCAACTTCGAAACCAAGAAATTTGGGCCACC	562	XX	Abbott CA, Gorrell MD;		
Db	584	TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr	603	XX	WPI; 2002-454646/48.		
				XX	N-PSDB; AAD38956.		
				XX	New dipeptidyl peptidase (Dpp) peptides, useful for screening inhibitors		
				XX	of DPP catalytic activity, which may be employed to treat e.g. neoplasia,		
				XX	type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV		
				XX	infection.		

PS Example; Fig 1; 91pp; English.

XX The present invention relates to dipeptidyl peptidase (DPP) proteins and polynucleotides encoding such proteins. The DPP peptides are useful for screening inhibitors of DPP catalytic activity. The inhibitors are useful for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV (human immunodeficiency virus) infection. The present sequence is human DPP8 protein

XX SQ Sequence 882 AA;

Alignment Scores:

Pred. No.: 1-75e-204 Length: 882

Score: 1836.50 Matches: 360

Percent Similarity: 67.5% Conservative: 0

Best Local Similarity: 67.5% Mismatches: 0

Query Match: 94.8% Indels: 173

DB: 5 Gaps: 1

US-10-825-632-8 (1-1083) x AA24170 (1-882)

QY 2 GAAGAGATGCCAGATCAGCTGGATCGCTACCTTTGTTCTCCAAAGAATTTGATAGA 61  
 DB 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263  
 QY 62 TATTCGGCTATGTGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTCCT 121  
 DB 264 TyrSerGlyTyrTrpCysProGlyAlaGluThrThrProSerGlyGlyLysIleLeu 283  
 QY 122 AGAATCTATATGAAGAAATGATGAATCTGAGTGGAAATATTTCATGTTACATCCCT 181  
 DB 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303  
 QY 182 ATGTGGAAACAAGAGGAGGAGATTCATTCGGTTATCCTAAACAGTACACCAATCCT 241  
 DB 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323  
 QY 242 AAAGTCACTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAGGATCATA 295  
 DB 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343  
 QY 295 ----- 295  
 DB 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363  
 QY 295 ----- 295  
 DB 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383  
 QY 295 ----- 295  
 DB 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403  
 QY 295 ----- 295  
 DB 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423  
 QY 295 ----- 295  
 DB 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443  
 QY 295 ----- 295  
 DB 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463  
 QY 295 ----- 295  
 DB 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483  
 QY 295 ----- 295  
 DB 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503

QY 296 -----? GTTGATGAAGTCAGAAAGCTGCTATAT 322  
 DB 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523  
 QY 323 TTGAAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTAGCTAGTCAGTTACGTAAAT 382  
 DB 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValIserTyrValAsn 543  
 QY 383 CCTGGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCAGTCAGTCAG 442  
 DB 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563  
 QY 443 CACTGTGACTTCTTTATAAGTATAGTAACACAGAGAAATCCACACTGTGTCCTCCCTT 502  
 DB 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValIserLeu 583  
 QY 503 TACAAGCTATCAAGTCCCTGACATGACCACTTGCACAAACAAAGGAATTTGGGCCACC 562  
 DB 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
 QY 563 ATTTTCGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGA 622  
 DB 604 IleLeuAspSerAlaGlyProLeuProAspLysTyrThrProGluIlePheSerPheGlu 623  
 QY 623 AGTACTACTGGAATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 682  
 DB 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
 QY 683 AAGAATATCTACTGTCTCTTTCATATATGGTGGTCTCTCAGGTGCAGTTGGTGAATAAT 742  
 DB 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663  
 QY 743 CGGTTTAAAGGAGTCAGATTTCCGTTGAATACCTACCTAGCTCTCTAGGTATGTGGTT 802  
 DB 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
 QY 803 GTAGTGATAGACAACAGGGATCTCTGTCACCGAGGCTTAAATTTCAAGGCGCTTTAAA 862  
 DB 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
 QY 863 TATAAATGGTCAAAATAGAAATTCAGATCAGGTGGAAGGACTCCCAATATCTAGTCTCT 922  
 DB 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
 QY 923 CGATATGATTTTCATTTAGCTTAGATCGTGGGATCCAGGCTGGTCTCTATCGAGGATAC 982  
 DB 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyTyr 743  
 QY 983 CTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGTTGCTATTCTCGGGCC 1042  
 DB 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
 QY 1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 1081  
 DB 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776  
 RESULT 7  
 AAU74749  
 ID AAU74749 standard; protein; 882 AA.  
 XX  
 AC AAU74749;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE Human protease PRTS-9 protein sequence.  
 XX  
 KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
 KW cell proliferative disorder; developmental disorder; epilepsy;  
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
 KW reproductive disorder; endometriosis.  
 XX





Score: 1836.50 Matches: 360  
Percent Similarity: 67.5% Conservative: 0  
Best Local Similarity: 67.5% Mismatches: 0  
Query Match: 94.8% Indels: 173  
DB: 5 Gaps: 1

US-10-825-632-8 (1-1083) x ADI17086 (1-882)

QY 2 GAAGAAGATGCACATCAGTCGAGTCGCTACCTTGTCTCCAAGAAGAAATTTGATAGA 61  
|||||  
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263  
|||||  
QY 62 TATTCTGGCTATTGCTGCTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTCCTT 121  
|||||  
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrProSerGlyGlyIleLeu 283  
|||||  
QY 122 AGAATTCATATGAAGAAATGATGATCTGAGGTGGAATTTATTCATGTTACATCCCT 181  
|||||  
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303  
|||||  
QY 182 ATGTTGGAACAGGAGGAGATTCATTCCTGTTATCTTAAACAGGTACAGCAATCCT 241  
|||||  
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323  
|||||  
QY 242 AAAGTCACTTTAAGATCTCAGAAATATGATGATGCTGAAGGAGGATCATA----- 295  
|||||  
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343  
|||||  
QY 295 ----- 295  
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363  
|||||  
QY 295 ----- 295  
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383  
|||||  
QY 295 ----- 295  
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403  
|||||  
QY 295 ----- 295  
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423  
|||||  
QY 295 ----- 295  
Db 424 GluGluThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443  
|||||  
QY 295 ----- 295  
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463  
|||||  
QY 295 ----- 295  
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483  
|||||  
QY 295 ----- 295  
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503  
|||||  
QY 296 -----GTTGATGAAGTCAGAGCTCGTATAT 322  
|||||  
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523  
|||||  
QY 323 TTTGAAGGACCAAGACTCCCTTTAGAGCATACCTGTAGTACGTAGTCAGTTAGTAAAT 382  
|||||  
Db 524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 543  
|||||  
QY 383 CTTGAGAGGTGACAGGCTGACTGACCGTGGCTACTCATTCTGCTGTCAGTCAGCAG 442  
|||||  
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563  
|||||  
QY 443 CACTGTGACTTCTTTATTAAGTAAGTATAGTAACCAAGAAGATCCACACTGTGTGCTT 502  
|||||

Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
|||||  
QY 503 TACAAGCTATCAAGTCTGAGTGCACCCAACTTGCACAAAGAAATTTGGCCACC 562  
|||||  
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
|||||  
QY 563 ATTTTGGATTCCAGCAGGTCCTCTTCCCTGACTACTACTCTCCAGAAATTTCTTTTAA 622  
|||||  
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
|||||  
QY 623 AGTACTACTGGATTACATTGTATGGAGTCTCTCAAGCCTCATGATCTACAGCCTGGA 682  
|||||  
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
|||||  
QY 683 AGAAATATCTACTGTGCTGTTTCATATATATGTTGGTCTCAGGTGAGTGGTGAATAT 742  
|||||  
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663  
|||||  
QY 743 CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTATGTGTT 802  
|||||  
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
|||||  
QY 803 GTAGTGTAGTAAACAGGGGATCTGTGCACGAGGCTTAAATTTGAAGCGCCTTTAAA 862  
|||||  
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
|||||  
QY 863 TATAAATGGGTCAAAATAGAAATTTGACGATCAGTGCAGTGCAGGCTCTCTATGAGGATAC 922  
|||||  
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
|||||  
QY 923 CGATATGATTTTCAATTTAGATTCGTGTGGGCATCCAGGCTGCTCTATGAGGATAC 982  
|||||  
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyTyr 743  
|||||  
QY 983 CTCCTCCCTGATGGCATTAAATGCAGATCAGATATCTTCAGGTTGCTATTCCTGGGCC 1042  
|||||  
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
|||||  
QY 1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081  
|||||  
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776  
|||||

RESULT 9  
ABU07720  
ID ABU07720 standard; protein; 882 AA.  
XX  
AC ABU07720;  
XX  
DT 19-MAY-2003 (first entry)  
XX  
DE Human serine protease HIPHUM46.  
XX  
KW Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;  
KW serine protease activity modulation; dipeptidyl peptidase activity;  
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;  
KW Alzheimer's disease; paraneuronal palsy; Huntington's disease;  
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;  
KW irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;  
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;  
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;  
KW multiple sclerosis.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Region : 259..260  
FT /note= "Paired glutamates of the beta propeller domain"  
FT Active-site 739  
FT /label= Catalytic\_serine\_residue  
FT Active-site 817  
FT /label= Catalytic\_aspartate\_residue  
FT Active-site 849

FT /label= Catalytic\_histidine\_residue  
PN GB2374869-A.  
XX  
XX 30-OCT-2002.  
XX  
XX 22-JAN-2002; 2002GB-00001404.  
XX  
XX 23-JAN-2001; 2001GB-00001760.  
XX (GLAX ) GLAXO GROUP LTD.  
XX  
XX Edbrooke MR, Lewis AP;  
XX WPI; 2003-150703/15.  
XX N-PSDB; ABX12255.  
XX  
XX Identifying modulators of serine protease activity useful for treating  
XX musculoskeletal diseases, by contacting cell expressing a novel serine  
XX protease polypeptide with a compound and monitoring serine protease  
XX activity.  
XX  
XX Claim 10; Page 26-29; 38pp; English.  
XX  
XX The invention relates to a method of identifying a substance that  
XX modulates serine protease activity, comprising contacting a cell, such as  
XX a neuronal cell, lung cell, intestinal cell or a cell infected with a  
XX virus, expressing a serine protease polypeptide (HIPHUM 46), or its  
XX variant having dipeptidyl peptidase activity, or a serine protease  
XX isolated from the cell with a test substance and monitoring for serine  
XX protease activity. The method is useful for identifying a substance that  
XX modulates serine protease activity. A modulator of the serine protease is  
XX useful in the manufacture of a medicament for treatment or prophylaxis of  
XX a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus  
XX infection, Alzheimer's disease, parasupranuclear palsy, myotonic  
XX dystrophy, Huntington's disease, or amyotrophic lateral sclerosis.  
XX Additional disease that may be treated using modulators of the serine  
XX protease include malabsorption syndromes, irritable bowel syndrome, lung  
XX disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,  
XX rectal polyps, small bowel tumours, colorectal tumours, anaemia,  
XX dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple  
XX sclerosis. The present sequence represents the amino acid sequence of the  
XX human serine protease HIPHUM46  
XX  
SQ Sequence 882 AA;  
  
Alignment Scores:  
Pred. No.: 1,75e-204 Length: 882  
Score: 1836.50 Matches: 360  
Percent Similarity: 67.5% Conservative: 0  
Best Local Similarity: 67.5% Mismatches: 0  
Query Match: 94.8% Indels: 173  
DB: 6 Gaps: 1  
  
US-10-825-632-8 (1-1083) x ABU07720 (1-882)  
  
QY 2 GAAGAGATCCGAGATCAGCTGAGTCGCTACCTTGTCTCCAGAGAAATTTGATAGA 61  
DB 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluGluPheAspArg 263  
  
QY 62 TATTCTGGCTATTGGTGGTCCAAAAGCTGAAAACAACTCCAGTGGTGGTAAATTCCT 121  
DB 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283  
  
QY 122 AGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCT 181  
DB 284 ArgLeuLeuTyrGluGluAsnAspGluSerGluValGluLeuLeuHisValThrSerPro 303  
  
QY 182 ATGTTGGAACAAAGGAGGGCAGATTTCATTCCTGTTATTCCTAAACAGAGGTACAGCAATCCT 241  
DB 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323  
  
QY 242 AAAGTCACCTTTAAGATGTCAGAAATAATGATTGATGCTGAGGAAGGATCATATA----- 295

|||||  
324 LysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIleLeuAspVal 343  
295 -----  
344 IleAspLysGluLeuLeuGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363  
295 -----  
364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383  
295 -----  
384 ThrArgLeuGlnIleValLeuLeuSerProGluLeuPheIleProValGluAspAspVal 403  
295 -----  
404 MetGluArgGlnArgLeuLeuGluSerValProAspSerValThrProLeuLeuIleTyr 423  
295 -----  
424 GluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPheProGlnSer 443  
295 -----  
444 HisGluGluGluLeuGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463  
295 -----  
464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483  
295 -----  
484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503  
296 ----- GTGATGAAGTCAGAGCTGGTATAT 322  
504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523  
323 TTTCAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAAT 382  
524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 543  
383 CCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTCCTGTGTCAGTCAGT 442  
544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563  
443 CACTGTGACTTCTTTATAAGTAAATAGTATAGTAAACCAAGAAATCCACACTGTGTCCCTT 502  
564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
503 TACAAGCTATCAAGTCTCGAAGATGACCCCACTTCGCAACAAAGAAATTTTGGGCCACC 562  
584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
563 ATTTTGGATTTCAGCAGGTCTCTCTCGACTATATCTCTCCAGAAATTTCTCTTTGAA 622  
604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
623 AGTACTACTCGATTACATTGTATGGGATGCTCTCAAGCCCTCATGATCTACAGCTGGA 682  
624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
683 AAGAAATATCTACTGTGCTGTTCATATATAGTGGTCTCAGGTGCAGTGGTGAATAT 742  
644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663  
743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCCTCTAGGTTATGTGGTT 802  
664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
803 GTAGTGATGACACCAAGGGGATCTCTGACCCAGGGCTTAAATTTGAAGGGCCCTTTAAA 862  
|||||

Db	684	ValValIleAspAsnArgGlySerCysHisArgGlyLeuIysPheGluGlyAlaPheIys	703
Qy	863	TATAAATGGTCAAATAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCTTCT	922
Db	704	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	723
Qy	923	CGATATGATTCATTGATTCCTAGATCCTGTGGGATCCAGGCTGCTCTATGGAGGATAC	982
Db	724	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr	743
Qy	983	CTCTCCCTGATGGCATTAAATGCAGAGCTCAGATATCTTCAGGGTTTGCTATTGCTGGGGCC	1042
Db	744	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763
Qy	1043	CCAGTCTACTGTGGATCTTCTTATGATACAGGATACACG	1081
Db	764	ProValThrLeuThrIlePheTyrAspThrGlyTyrThr	776
RESULT 10			
AD117085			
ID	AD117085	standard; protein; 883 AA.	
XX	AC	AD117085;	
XX	XX		
DT	15-APR-2004	(first entry)	
DE	Murine NOVX protein homologue SeqID 621.		
XX	KW	mouse; murine; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;	
XX	KW	inflammation; autoimmune disorder; allergy; blood disorder;	
KW	KW	acquired immunodeficiency syndrome; AIDS; obesity; asthma;	
KW	KW	immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;	
KW	KW	Alzheimer's disease; infect.	
XX	OS	Mus musculus.	
XX	XX		
FN	WO200268649-A2.		
XX	PD	06-SEP-2002.	
XX	XX		
PF	31-JAN-2002;	2002WO-US002785.	
XX	31-JAN-2001;	2001US-0265395P.	
PR	31-JAN-2001;	2001US-0265412P.	
PR	31-JAN-2001;	2001US-0265514P.	
PR	31-JAN-2001;	2001US-0265517P.	
PR	02-FEB-2001;	2001US-0266406P.	
PR	05-FEB-2001;	2001US-0266767P.	
PR	07-FEB-2001;	2001US-0266975P.	
PR	08-FEB-2001;	2001US-0267459P.	
PR	09-FEB-2001;	2001US-0267823P.	
PR	15-FEB-2001;	2001US-0268974P.	
PR	26-FEB-2001;	2001US-0271664P.	
PR	27-FEB-2001;	2001US-0271839P.	
PR	27-FEB-2001;	2001US-0271855P.	
PR	02-MAR-2001;	2001US-0272788P.	
PR	02-MAR-2001;	2001US-0273046P.	
PR	14-MAR-2001;	2001US-0275925P.	
PR	14-MAR-2001;	2001US-0275947P.	
PR	14-MAR-2001;	2001US-0275950P.	
PR	14-MAR-2001;	2001US-0275989P.	
PR	15-MAR-2001;	2001US-0276448P.	
PR	15-MAR-2001;	2001US-0276450P.	
PR	16-MAR-2001;	2001US-0276397P.	
PR	16-MAR-2001;	2001US-0276768P.	
PR	20-MAR-2001;	2001US-0278652P.	
PR	26-MAR-2001;	2001US-0278775P.	
PR	26-MAR-2001;	2001US-0278778P.	
PR	29-MAR-2001;	2001US-0279882P.	
PR	29-MAR-2001;	2001US-0279884P.	
PR	30-MAR-2001;	2001US-0280147P.	
PR	11-APR-2001;	2001US-0282992P.	

PR	11-APR-2001;	2001US-0283083P.	
PR	20-APR-2001;	2001US-0285133P.	
PR	23-APR-2001;	2001US-0285749P.	
PR	03-MAY-2001;	2001US-0288327P.	
PR	03-MAY-2001;	2001US-0288504P.	
PR	29-MAY-2001;	2001US-0294047P.	
PR	30-MAY-2001;	2001US-0294473P.	
PR	08-JUN-2001;	2001US-0296964P.	
PR	18-JUN-2001;	2001US-0298959P.	
PR	19-JUN-2001;	2001US-0299324P.	
PR	13-AUG-2001;	2001US-0312020P.	
PR	16-AUG-2001;	2001US-0312889P.	
PR	21-AUG-2001;	2001US-0313390P.	
PR	28-AUG-2001;	2001US-0315470P.	
PR	31-AUG-2001;	2001US-0316447P.	
PR	07-SEP-2001;	2001US-0318115P.	
PR	07-SEP-2001;	2001US-0318118P.	
PR	12-SEP-2001;	2001US-0318740P.	
PR	19-SEP-2001;	2001US-0323379P.	
PR	18-OCT-2001;	2001US-0330245P.	
PR	18-OCT-2001;	2001US-0330308P.	
XX	14-NOV-2001;	2001US-0332701P.	
XX			
PA	(CURA-) CURAGEN CORP.		
XX			
PI	Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;		
PI	Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;		
PI	Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;		
PI	Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;		
XX			
DR	WPI; 2002-706998/76.		
XX			
PT	New NOVX polypeptides and nucleic acids, useful for preventing or		
PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,		
PT	atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or		
PT	pharmacogenomics.		
XX			
PS	Disclosure; SEQ ID NO 621; 1498pp; English.		
XX			
CC	This invention relates to a novel nucleic acids, and encoded polypeptides		
CC	thereof, which have properties related to the stimulation of biochemical		
CC	or physiological responses in a cell, tissue, organ or organism.		
CC	Specifically, it refers to the use of biologically active fragments for		
CC	diagnostic and prognostic assays and furthermore in the treatment of		
CC	diverse pathological conditions. The present invention describes novel		
CC	human and murine NOVX proteins, as well as methods to modulate their		
CC	expression using antisense oligos, ribozymes and peptide nucleic acids.		
CC	The NOVX polypeptides, polynucleotides and antibodies are useful in		
CC	treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,		
CC	atherosclerosis, cancer and diabetes. Furthermore, they may be used in		
CC	treating or preventing diseases such as inflammation, autoimmune		
CC	disorders, allergies, blood disorders, acquired immunodeficiency syndrome		
CC	(AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,		
CC	arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy		
CC	and epilepsy. Accordingly, these molecules have many activities including		
CC	cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,		
CC	hemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,		
CC	antiasthmatic, nephrotropic, antiarthritic, hepatotropic,		
CC	neuroprotective, nootropic, antibacterial, virucide, antiparasitic,		
CC	relaxant and anticonvulsant. In addition, they are useful in screening		
CC	assays to identify small molecules that modulate or inhibit, for example,		
CC	neurogenesis, wound healing and angiogenesis. The nucleic acids are also		
CC	used as in chromosome mapping, tissue typing, preventive medicine and		
CC	pharmacogenomics. This polypeptide is a homologue of a human NOVX protein		
CC	of the invention.		
XX			
SQ	Sequence 883 AA;		
Alignment Scores:			
Pred. No.:	3,85e-195	Length:	883
Score:	1756.50	Matches:	343
Percent Similarity:	65.7%	Conservative:	7



Best Local Similarity: 64.4%		Matches: 10	
Query Match: 90.6%		Indels: 173	
DB: 5		Gaps: 1	
US-10-825-632-8 (1-1083) x AD117085 (1-883)			
QY	2	GAGAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGA	61
DB	245	GlulGluAaspProArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAaspArg	264
QY	62	TATCTGGCTATTGGTGTCCAAAGCTGAAACAACTCCACAGTGGTGGTAAATTCCT	121
DB	265	TyrSerGlyTyrTrpTrpCysProGlnAlaGluArgThrProSerGlyGlyLysIleLeu	284
QY	122	AGAATTCATATCAAGAAATCATCATCTGAGTGGGAAATTTATTCATGTTACATCCCT	181
DB	285	ArgIleLeuTyrGluGluAenAaspGluSerGluValGluIleHisValThrSerPro	304
QY	182	ATGTTGGAACAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACAGCAATCCT	241
DB	305	MetLeuGluThrArgArgAlaAaspSerPheArgTyrProLysThrGlyThrAlaAenPro	324
QY	242	AAAGTCACATTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATCATA	295
DB	325	LysValThrPheLysMetSerGluIleValValAaspAlaAlaGlyGlyIleAaspVal	344
QY	295	-----	295
DB	345	IleAaspLysGluLeuValGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla	364
QY	295	-----	295
DB	365	ArgAlaGlyTrpTrpProGluGlyLysHisAlaTrpSerIleLeuLeuAaspArgSerGln	384
QY	295	-----	295
DB	385	ThrHisLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAaspAla	404
QY	295	-----	295
DB	405	MetAaspArgGlnArgLeuIleGluSerValProAaspSerValThrProLeuIleIleTyr	424
QY	295	-----	295
DB	425	GlulGluThrThrAaspIleTrpIleAenIleHisAaspIlePheHisValPheProGlnThr	444
QY	295	-----	295
DB	445	HisGluAaspGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu	464
QY	295	-----	295
DB	465	TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro	484
QY	295	-----	295
DB	485	AlaProSerAaspPheLysCysProIleLysGluGluIleThrIleThrSerGlyGluTrp	504
QY	296	-----GTTGATGAAGTCAGAGCTGGTATAT	322
DB	505	GluValLeuGlyArgHisGlySerAenIleTrpValAaspGluAlaArgLysLeuValTyr	524
QY	323	TTTGAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAAT	382
DB	525	PheGluGlyThrLysAaspSerProLeuGluHisHisLeuTyrValThrSerTyrAlaAen	544
QY	383	CCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCCTGCTGCATCAGTCAG	442
DB	545	ProGlyGluValValArgLeuThrAaspArgGlyTyrSerHisSerCysLysLeuSerArg	564
QY	443	CACGTGACTCTTTTATAAGTAGATAGTAACCAAGAAAGATCCACACTGTGTCTCCCTT	502
DB	565	HisCysAaspPhePheIleSerLysTyrSerAenGlnLysAenProHisCysValSerLeu	584
QY	503	TACAGCTATCAAGTCTCTGAGATGACCCAACTTCTGAAAACAAGAGATTTTGGCCACC	562
DB	585	TyLysLeuSerSerProGluAaspProValHisLysThrLysGluPheTrpAlaThr	604
QY	563	ATTTTGATTACAGAGGTCTCTCTCTGACTATATCTCTCCAGAAAATTTCTCTTTGAA	622
DB	605	IleLeuAaspSerAlaGlyProLeuProAaspTyrThrProProGluIlePheSerPheGlu	624
QY	623	AGTACTACTGGATTTACATTGTATGGGATGCTCTACAAGCCCTCATGATCTACAGCCTGGA	682
DB	625	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAaspLeuGlnProGly	644
QY	683	AAGAAATATCTACTGTCTGTTTCATATATGTTGGTCTCTCAGGTGCAGTTGGTGAATAAT	742
DB	645	LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAenAen	664
QY	743	CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTATGTGTT	802
DB	665	ArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAaserLeuGlyTyrValVal	684
QY	803	GTAGTGATACACAAACAGGGATCCGTGACCCAGGGCTTAAATTTGAAGGCGCTTTAAA	862
DB	685	ValValIleAaspAenArgGlySerCysHisAaGlyLeuLysPheGluGlyAlaPheLys	704
QY	863	TATAAATGGTCCAAATAGAAATTCAGATCAGGTGGAAGGACTCCATATCTAGCTCT	922
DB	705	TyLysMetGlyGlnIleGluIleAaspGlnValGluGlyLeuGlnTyrLeuAaser	724
QY	923	CGATATGATTCATTCAGTACTAGATCGTGTGGCATCCAGGCTGGTCTATCGGATAC	982
DB	725	GlnTyrAaspPheIleAaspLeuAaspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr	744
QY	983	CTCTCCCTGATGGCATTAAATGCAGAGTCAAGATATCTTCAGGTTGCTATTGCTGGGGCC	1042
DB	745	LeuSerLeuMetAlaLeuMetGlnArgSerAaspIlePheArgValAlaIleAlaGlyAla	764
QY	1043	CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081	
DB	765	ProValThrLeuTrpIlePheTyrAaspThrGlyTyrThr 777	
RESULT 11			
ABU92026			
ID	ABU92026	standard; protein; 831 AA.	
XX	AC	ABU92026;	
XX	DT	15-JUL-2003 (first entry)	
XX	DE	Human protein modification and maintenance molecule-6 (PMMW-6).	
XX	KW	Human; protein modification and maintenance molecule; PMMW; cancer;	
XX	KW	cell proliferation disorder; atherosclerosis; neurological disorder;	
XX	KW	epilepsy; Huntington's disease; stroke; immune disorder; allergy;	
XX	KW	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;	
XX	KW	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;	
XX	KW	infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;	
XX	KW	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;	
XX	OS	antiflammatory; thymomimetic.	
XX	PN	Homo sapiens.	
XX	PN	WO2003031939-A2.	
XX	PD	17-APR-2003.	
XX	PF	11-OCT-2002; 2002WO-US032850.	
XX	PR	12-OCT-2001; 2001US-0329689P.	
XX	PR	25-OCT-2001; 2001US-0335703P.	
XX	PR	09-NOV-2001; 2001US-034887P.	
XX	PR	28-NOV-2001; 2001US-0334145P.	
XX	PR	06-DEC-2001; 2001US-0337451P.	

PR 14-DEC-2001; 2001US-0340584P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;  
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;  
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;  
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
XX  
DR WPI; 2003-430274/40.  
DR N-PSDB; ACA92421.

XX New human protein modification and maintenance molecules (PMMM), useful  
PT for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or  
PT infections.

XX Claim 1; Page 242-243; 31pp; English.

XX The present invention relates to the isolation of human protein  
CC modification and maintenance molecules (PMMM), and the polynucleotide  
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM  
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful  
CC for diagnosing a condition or disease associated with the expression of  
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and  
CC generating an expression profile of a sample containing the  
CC polynucleotides. The diseases or conditions associated with decreased  
CC expression or overexpression of PMMM are cell proliferation disorders  
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,  
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
CC syndrome), gastrointestinal or epithelial disorders, and infections. The  
CC PMMM polypeptides or their fragments are useful in screening compounds  
CC for effectiveness as agonists or antagonists of the polypeptides, or in  
CC altering the expression of the target polynucleotide and compounds that  
CC specifically bind to, or modulate the activity of the polypeptide.  
CC ABU92021-ABU92060 represent the human PMMM polypeptides of the invention

XX Sequence 831 AA;

Alignment Scores:  
Pred. No.: 3 616-169 Length: 831  
Score: 1534.00 Matches: 309  
Percent Similarity: 58.0% Conservative: 0  
Best Local Similarity: 58.0% Mismatches: 0  
Query Match: 79.2% Indels: 224  
DB: 6 Gaps: 2

US-10-825-632-8 (1-1083) x ABU92026 (1-831)

QY 2 GAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAATTGTGATGA 61  
DB 244 GLUGLUSPALARGSERALAGLYVALLALATHRPHEVALLEUGLUGLUPHEASPARG 263  
QY 62 TAATCTGGCTATTGGTGTGTCCTCAAGCTGAAACCACTCCAGCTGGTGTAAATTCCT 121  
DB 264 TYR-SERGLYTYRTRP-CYSPOLYSALAAGLUTHRTHRPRO-SERGLYGLYLYSLEU 283  
QY 122 AGAATCTATATGAAGAAATGATGAAATCTGAGGTGGAAATATTATGTTACATCCCT 181  
DB 284 ARGILEUITYRGLUGLUSASPGLU-SERGLUVALGLULLEIHISVALTHR-SERPRO 303  
QY 182 ATGTTGGAACACAGGAGGCAGATTCATCCGTTATCTCTAAACAGAGTACGCAATCCT 241  
DB 304 METLEUGLUTHRARGALASPSP-PHEARGTYRPROLYSTHRGLYTHR-ALASPRO 323  
QY 242 AAAGTCACATTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATCATATA 295  
DB 324 LYSVALTHR-PHELYSMETSERGLULLEMETILEASPALAGLUGLYARGILEILEAS 343  
QY 295 ----- 295

DB 344 ILEASPLYSGLULEULEGPNPRO-PHEGLULLEULEPHEGLUGLYVALGLUTYRILEALA 363  
QY 295 ----- 295  
DB 364 ARGALAGLYTRP-THRPROGLUGLYLSTYRALA-TRP-SERILELEULEUASPARGSERGLN 383  
QY 295 ----- 295  
DB 384 THRARGLEUGLNILEVALLEULESERPROGLUGLUPHEILEPROVALGLUASPVAL 403  
QY 295 ----- 295  
DB 404 METGLUARGLNARGLEULEGLUSERVALPROASPSERVALTHRPROLEULEILETYR 423  
QY 295 ----- 295  
DB 424 GLUGLUTHRTHRASPILETRP-ILEASNILEHISASPILEPHEHISVALPHEPROGLN-SER 443  
QY 295 ----- 295  
DB 444 HISGLUGLUGLUGLUPHEILEPHEALA-SERGLUCYLSYSTRGLYPHEARGHISLEU 463  
QY 295 ----- 295  
DB 464 TYR-LSILETHR-SERILELEULYSGLUSERLYSTYR-LSY-ARG-SER-SERGLYGLY-LEUPRO 483  
QY 295 ----- 295  
DB 484 ALA-PRO-SER-ASP-PHELYSCYSPRO-ILEYSGLU-GLULLE-ALA-ILE-THR-SER-GLY-GLU-TRP 503  
QY 296 -----GTTGATGAAGTCAGAGCTGGTATAT 322  
DB 504 GLU-VALLEUGLYARGHISGLYSERASNILEGLN-VALASPGLU-VALARGLEU-VALTYR 523  
QY 323 TTTGAGGACCAACAGCTCCCTTTAGAGCATACCTGTAGTAGTCAGTTACGTAAAT 382  
DB 524 PHEGLUGLYTHR-LSY-ASP-SER-PROLEUGLULHISLEU-VALSERLYR-VALSERLYR-VALASN 543  
QY 383 CCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTCCTTGTGTCAGTCAG 442  
DB 544 PROGLUGLUTHRARGLEUTHRASPARGLYTYR-SERHIS-SER-CYS-ILE-SERGLN 563  
QY 443 CACTGTGACTTCTTTATAAGTAACTAGTATAACCAAGAGAATCCACACTGTGTGTCCTT 502  
DB 564 HIS-CYS-ASP-PHE-ILE-SERLYSTYR-SERASNGLNLYSASNPROHIS-CYS-VALSERLEU 583  
QY 503 TACAAGCTATCAAGTCTCTGAAAGATGCCCACTTGCAAAACAAAGGAATTTTGGCCACC 562  
DB 584 TYR-LSY-LEU-SER-SER-PROGLU-ASP-ASP-PRO-THR-CYS-LSY-THR-LSY-GLU-PHE-TRP-ALA-THR 603  
QY 563 ATTTTGGATTCACAGAGTCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAA 622  
DB 604 ILELEUASP-SER-ALA-GLY-PROLEU-PROASP-TYR-THR-PROGLULLE-PHE-SER-PHEGLU 623  
QY 623 AGTACTACTGGATTACATTTGATGGATGCTCTCAAGCTCTCATGATCTACAGCTCGA 682  
DB 624 SER-THR-THR-GLY-PHE-THR-LEU-TYR-GLY-MET-LEU-TYR-LSY-PROHIS-ASP-LEU-PROGLY 643  
QY 683 AAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGTGCAGTGTGTCGTAATAAT 742  
DB 644 LYSLYSTYR-PRO-THR-VALLEU-PHE-ILE-TYR-GLY-GLY-PROGLIN-VALGLINLEU-VALASN-ASN 663  
QY 743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTCTAGGTATGTGGTT 802  
DB 664 ARG-PHELYSGLYVALLYSTYR-PHE-ARG-LEU-ASN-THR-LEU-ALA-SER-LEU-GLYTYR-VAL 683  
QY 803 GTAGTGATGACCAACAGGGGATCTCTGTCCAGGAGGCTTAAATTTGAAGGCGCCTTTAAA 862  
DB 684 VAL-VAL-ILE-ASP-ASN-ARGGLY-SER-CYS-HIS-ARG-GLY-LEU-LSY-PHE-GLU-GLY-ALA-PHELYS 703  
QY 863 TATTAATGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCATATATCTAGCTTCT 922  
DB 704 TYR-LSY-MET----- 706

QY 923 CGATATGATTTCATTGACTTAGATCGTGTGGCATCCAGCGCTGGTCTCTATGGAGGATAC 982  
Db 706 ----- 706  
QY 983 CTCCTCCGTGATGCATTATGATGAGAGGTGAGATATCTTCAGGGTGTCTATGCTGGGGCC 1042  
Db 707 -----ValAlaIleAlaGlyAla 712  
QY 1043 CGAGTCACTCTGGGATCTTCATGATACAGATACAG 1081  
Db 713 ProvalThrLeuTrpIlePheTyrAspThrGlyTyrThr 725  
RESULT 12  
ADE78977  
ID ADE78977 standard; protein; 493 AA.  
AC ADE78977;  
XX  
XX  
XX 29-JAN-2004 (first entry)  
XX  
XX Human protein modification and maintenance molecule (PMMM) -15.  
XX  
XX protein modification and maintenance molecule; PMMM;  
KW protein modification; protein maintenance; protein function;  
KW protein conformation; protein stabilisation; protein degradation; kinase;  
KW phosphatase; protease; protease inhibitor; isomerase; transferase;  
KW molecular chaperone; anti-HIV; anti-allergic; anti-inflammatory;  
KW melanaemic; antiparkinsonian; nontropic; anticonvulsant;  
KW antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;  
KW cytotatic; hepatotropic; dermatological; antidiabetic; nephrotropic;  
KW antitumor; thyromimetic; neuroprotective; osteopathic; antidiabetic;  
KW antiparasitic; antihelminthic; antipsoriatic; uropathic; ophthalmological;  
KW antirheumatic; haemostatic; antibacterial; virucide; protozoacide;  
KW fungicide; gene therapy; cell proliferative disorder; arteriosclerosis;  
KW hepatitis; polychaemia vera; psoriasis; primary thrombocytopaenia;  
KW cancer; developmental disorder; anaemia; mental retardation;  
KW neurological disorder; Alzheimer's disease; Parkinson's disease;  
KW epilepsy; autoimmune disorder; inflammatory disorder; AIDS; allergies;  
KW asthma; autoimmune thyroiditis; Crohn's disease; diabetes mellitus;  
KW glomerulonephritis; Goodpasture's syndrome; multiple sclerosis;  
KW arthritis; osteoporosis; pancreatitis; Sjogren's syndrome;  
KW microbial infection; human.  
XX  
XX Homo sapiens.  
XX  
XX WO2003063688-A2.  
XX  
XX 07-AUG-2003.  
XX  
XX 23-JAN-2003; 2003WO-US002500.  
XX  
XX 25-JAN-2002; 2002US-0351928P.  
XX  
XX 25-FEB-2002; 2002US-0359903P.  
XX  
XX 21-MAR-2002; 2002US-0366837P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Hafalia AJA; Li JX, Gervad AE, Chawla NK, Sprague WW, Lee SY;  
PI Chang H, Elliott VS, Ramkumar J, Khare R, Emerling BM, Kable AE;  
PI Tang YT, Yue H, P, Gietzen KJ, Lee S, Swarnakar A, Baughn MK;  
PI Wilson AD, Jin P, Chien D, Hawkins PR, Jiang X, Jackson AA;  
PI Bhatia U, Burrill JB, Blake JJ, Ho A, Zheng W, Ison CH, Marquis JP;  
PI Iran UK, Lal PG, Warren BA, Xu Y, Honchell CD, Becha SD;  
PI Lehr-Mason FM;  
XX  
XX WPI; 2003-636761/60.  
XX  
XX N-PSDB; ADE79035.  
XX  
XX New human protein modification and maintenance molecules and  
PT polynucleotides, useful for diagnosing, treating or preventing autoimmune  
PT or inflammatory disorders (e.g. AIDS, allergy or anemia), multiple  
PT sclerosis or cancer.

XX  
PS  
XX  
CC  
CC This invention relates to novel isolated human proteins, which are human  
CC protein modification and maintenance molecules (PMMM). The cellular  
CC processes regulating modification and maintenance of protein molecules  
CC coordinate their function, conformation, stabilisation and degradation.  
CC Each of these processes is mediated by key enzymes or proteins such as  
CC kinases, phosphatases, proteases, protease inhibitors, isomerases,  
CC transferases and molecular chaperones. Compounds which modulate the  
CC proteins of the invention may have anti-HIV, antiallergic,  
CC antiinflammatory, antianaemic, antiparkinsonian, nontropic,  
CC anticonvulsant, antiarteriosclerotic, antiasthmatic, immunosuppressive,  
CC antithyroid, cytotatic, hepatotropic, dermatological, antidiabetic,  
CC nephrotropic, antitumor, thyromimetic, neuroprotective, osteopathic,  
CC antirheumatic, antihelminthic, antipsoriatic, uropathic, ophthalmological,  
CC fungicide or gene therapy activities. The DNA sequence which encodes the  
CC protozoacide or fungicide activities. The DNA sequence which encodes the  
CC proteins of the invention may be useful for gene therapy. The human  
CC protein modification and maintenance molecules (PMMM), the DNA sequences  
CC which encode them and their modulating compounds are useful for  
CC diagnosing, treating or preventing disorders associated with aberrant  
CC expression of PMMM, particularly cell proliferative disorders (for  
CC example arteriosclerosis, hepatitis, polycythaemia vera, psoriasis,  
CC primary thrombocytopaenia or cancer), developmental disorders (for  
CC example anaemia or mental retardation), neurological disorders (for  
CC example Alzheimer's disease, Parkinson's disease or epilepsy),  
CC autoimmune/inflammatory disorders (for example AIDS, allergies, asthma,  
CC autoimmune thyroiditis, Crohn's disease, diabetes mellitus,  
CC glomerulonephritis, Goodpasture's syndrome, multiple sclerosis,  
CC arthritis, osteoporosis, pancreatitis, Sjogren's syndrome) or microbial  
CC infections. The present sequence is the amino acid sequence of a human  
XX PMMM of the invention.  
SQ Sequence 493 AA;  
Alignment Scores:  
Pred. No.: 3,31e-169 Length: 493  
Score: 1533.50 Matches: 305  
Percent Similarity: 69.6% Conservative: 0  
Best Local Similarity: 69.6% Mismatches: 2  
Query Match: 79.1% Indels: 132  
DB: 7 Gaps: 2  
US-10-825-632-8 (1-1083) x ADE78977 (1-493)  
QY 9 ATGCCAGATCAGCTGGAGTCGCTGCTTTCTTCCAAAGAAATTTGATAGATATTCTG 68  
Db 1 MetProAspGlnLeuGluSerLeuProLeuPheSerLysLysAsnLeuIleAspIleLeu 20  
QY 69 GCTATTGGTGGTCCAAAAGCTGAACAACTCCAGTGGTGGTAAATTCCTTAGAATTC 128  
Db 21 AlaIleGlyGlyValGlnLysLeuLysGlnLeuProValValValLysPheLeuGluPhe 40  
QY 129 TATATGAAGAAATGATGATCTGAGTGGGAATTTATTCATCTACATCCCTATGTTGG 188  
Db 41 TyrMetLysLysMetMetAsnLeuArgTrpLysLeuPheMetLeuHisProLeuLysTrp 60  
QY 189 AAACAGAGGAGG-GCAGATTTCATTCCTGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTC 247  
Db 61 LysGlnGlyArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnProLysVal 80  
QY 248 ACTTTTAAGATGTCGAAATAATATGATGCTGTAAGGAAGG----- 289  
Db 81 ThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleHisAspIlePheHis 100  
QY 289 ----- 289  
Db 101 ValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThr 120  
QY 289 ----- 289  
Db 121 GlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSer 140

QY 289 ----- 289  
Db 141 SerGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluLeAlaIle 160  
QY 290 -----ATCATAGTTGATGAAGTC 307  
Db 161 ThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluVal 180  
QY 308 AGAAGCTGGTATATTTGAAGGACACCAAGATCCCTTTAGAGATCACCTGTACGTA 367  
Db 181 ArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrVal 200  
QY 368 GTCAGTTAGTAAATCCTGGAGAGCTGACAAAGGCTGACCGTGGCTACTCACATCT 427  
Db 201 ValSerTyrValAsnProGlyValThrArgLeuThrAspArgGlyTyrSerHisSer 220  
QY 428 TGCTGCATCAGTCAGCACTGTGACTTCTTTATAAGTATAGTAAACCAAGAAATCCA 487  
Db 221 CysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnPro 240  
QY 488 CACTGTGTCTCTTTACAGCTATCAAGTATCAAGTACCACTTGCACCAACAAG 547  
Db 241 HisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLys 260  
QY 548 GAATTTTGGCCACCATTTTGGATTTCAGCAGTCTCTCTGCTGACTATATCTCTCCAGAA 607  
Db 261 GluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGlu 280  
QY 608 ATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCTCAT 667  
Db 281 IlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHis 300  
QY 668 GAICTACAGCTGGAAAGAAATATCTACTGTGCTGTTCATATATGTGTCTCTCAGTG 727  
Db 301 AspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnVal 320  
QY 728 CACTGTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCT 787  
Db 321 GlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSer 340  
QY 788 CTAGGTATGTGTTGTTAGTATAGACACAGGGATCTCTCTCACGAGGCTTAAATTT 847  
Db 341 LeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPhe 360  
QY 848 GAAGGCGCTTTAAATATATAATGGTCAATAGAAATGACATGAGTGGAGGAGCTC 907  
Db 361 GluGlyAlaPheLysTyrLysMet----- 368  
QY 908 CAATATCTAGCTTCTCGATATGATTTTCATTGACTTTAGATCGTGTGGCATCCACGGCTGG 967  
Db 368 ----- 368  
QY 968 TCCTATGGAGGATACCTCTCCCTGATGCGATTAAATGCAGAGGTGAGATATCTTCAGGGTT 1027  
Db 369 -----Val 369  
QY 1028 GCTATTGCTGGGCGCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081  
Db 370 AlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 387

RESULT 13

AAB93565  
ID AAB93565 standard; protein; 632 AA.  
XX  
AC AAB93565;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:12964.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.  
XX  
EN EPI074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
XX 28-JUL-2000; 2000EP-00116126.  
XX  
XX 29-JUL-1999; 99JP-00248036.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-00118776.  
PR 02-MAY-2000; 2000JP-00183767.  
PR 09-JUN-2000; 2000JP-00241899.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
XX WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
XX  
XX Claim 8; SEQ ID NO 12964; 2537pp + Sequence Listing; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the polynucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention  
XX  
SQ Sequence 632 AA;  
  
Alignment Scores:  
Pred. No.: 1,61e-168 Length: 632  
Score: 1528.00 Matches: 308  
Percent Similarity: 57.8% Conservative: 0  
Best Local Similarity: 57.8% Mismatches: 1  
Query Match: 78.8% Indels: 224  
DB: 4 Gaps: 2  
  
US-10-825-632-8 (1-1083) x AAB93565 (1-632)  
QY 2 GAAGAAGATGCCAGATCAGCTGAGTGCCTTGTCTTCTCAAGAAGAATTTGATAGA 61  
Db 45 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 64  
QY 62 TATTCTGCTATTGGTGTCTCCAAAGCTGAAACACATCCAGCTGGGTAAATTTCT 121  
Db 65 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 84  
QY 122 AGAATTCATATGAGAAATGATGATCTGAGGTGGAATATTTCATGTTACATCCCT 181

Db 85 ArgIleLeuTyrGluGluAsnAepGluSerGluValGluIleIleHisValThrSerPro 104  
QY 182 ATGTTGGAACAAAGGAGGCGAGATTTCATTCCTGTTATCTTAAACACAGGTACAGCAATCCT 241  
Db 105 MetLeuGluThrArgAlaAepSerPheA-gTyrProLysThrGlyThrAlaAsnPro 134  
QY 242 AAAGTCACCTTTTAAAGATGTCAGAAATAAATGATTGATGCTGAAGNAGATCATATA- 295  
Db 125 LysValThrPheLysMetSerGluIleMetIleAepAlaGluGlyArgIleAepVal 144  
QY 295 ----- 295  
Db 145 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 164  
QY 295 ----- 295  
Db 165 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAepArgSerGln 194  
QY 295 ----- 295  
Db 185 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAepAspVal 204  
QY 295 ----- 295  
Db 205 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 224  
QY 295 ----- 295  
Db 225 GluGluThrThrAspIleTrpIleAsnIleHisAepIlePheHisValPheProGlnSer 244  
QY 295 ----- 295  
Db 245 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 264  
QY 295 ----- 295  
Db 265 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 284  
QY 295 ----- 295  
Db 285 AlaProSerAspPheLysCysProIleLysGluGluLeuAlaIleThrSerGlyGluTrp 304  
QY 296 -----GTTGATGAAGTCAGAGCGCTGATAT 322  
Db 305 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 324  
QY 323 TTGGAAGGCCAACAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAAT 382  
Db 325 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 344  
QY 383 CTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTCCTGCTGCATCAGTCAG 442  
Db 345 ProGlyGluValThrArgLeuThrAspArgSerTyrSerHisSerCysLysSerGln 364  
QY 443 CACTGTGACTCTTTTATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCCCTT 502  
Db 365 HisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 384  
QY 503 TACAAGCTATCAAGTCCTGAAGATGACCCCAACTTGCACAAACAAAGAAATTTGGGCCACC 562  
Db 385 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 404  
QY 563 ATTTTGGATTGACAGAGGTCCTCTTCCTGACTATACCTCCAGAAATTTCTTTGAA 622  
Db 405 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 424  
QY 623 AGTACTACTGATTTTACATTTGATGGGATGCTTACAGCCTCATGATCTACAGCCTGGA 682  
Db 425 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 444  
QY 683 AGAAATATCTTACTGTGCTGCTTCATATATGTTGGTCTCAGGTGAGTTGGTGAATAAT 742  
Db 445 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 464

QY 743 CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTTATGTGTT 802  
Db 465 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 484  
QY 803 GTACTGTATAGACAACAGGGGATCCTGTCCACCGAGGCTTTAAATTTGAAGGGCGCTTTAAA 862  
Db 485 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 504  
QY 863 TATAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922  
Db 505 TyrLysMet----- 507  
QY 923 CGATATGATTTTCATTGACTTTAGATCGTGTGGCATCCACGGCTGCTCTATGAGGATAC 982  
Db 507 ----- 507  
QY 983 CTCTCCCTGATGCATTAAATGCAGAGTGCAGATATCTTCAGGGTTGCTATTGCTGGGGCC 1042  
Db 508 -----ValAlaIleAlaGlyAla 513  
QY 1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081  
Db 514 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 526  
RESULT 14  
ADR41398  
ID ADR41398 standard; protein; 587 AA.  
XX  
AC ADR41398;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Human CD-like molecule HDQES11, SEQ ID NO:197.  
XX  
KW Human; CD-like molecule; cluster of differentiation; diagnosis;  
KW prevention; immune disorder; immunodeficiency; autoimmune disorder;  
KW blood-related disorder; haematological disorder; haemostatic disorder;  
KW thrombolytic disorder; hyperproliferative disorder; cancer; tumour;  
KW apoptotic disorder; cardiovascular disorder; respiratory disorder;  
KW angiogenic disorder; neovascularisation; neurological disorder;  
KW endocrine disorder; reproductive system disorder; infectious disease;  
KW gastrointestinal disorder; drug screening; tissue regeneration;  
KW chemotaxis; gene therapy; antibody therapy; drug targeting;  
KW chromosome mapping; forensic analysis; immunophenotyping; cytostatic;  
KW haemostatic; tranquiliser; vulnery; antiinflammatory; nephrotropic;  
KW cardiant; anti-allergic; anti-HIV; antirheumatic; antiarthritic;  
KW antipsoriatic; immunosuppressive; vasotropic; nootropic;  
KW antithyroid; thyromimetic; gynaecological; viucide; hepatotropic;  
KW antibacterial; dermatological; chromosome 15q22.  
OS Homo sapiens.  
XX  
PN WO200226930-A2.  
XX  
PD 04-APR-2002.  
XX  
PF 25-SEP-2001; 2001WO-US029838.  
XX  
PR 26-SEP-2000; 2000US-0235484P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Birse CE;  
XX  
DR WPI; 2002-405050/43.  
DR N-PSDB; ADR41222.  
XX  
PT Novel polynucleotides and polypeptides useful for treating, preventing or  
PT ameliorating cardiovascular, renal, neurovascular, and autoimmune  
PT disorders.  
XX  
PS Claim 11; SEQ ID NO 197; 1243pp; English.

XX The invention relates to 167 novel human CD (cluster of differentiation)-  
 CC like molecules (ADR41388-ADR41563) and to cDNAs encoding them (seqid:11)-  
 XX  
 SQ Sequence 587 AA;

## Alignment Scores:

Pred. No.: 3,06e-168 Length: 587  
 Score: 1525.50 Matches: 304  
 Percent Similarity: 63.5% Conservative: 0  
 Best Local Similarity: 63.5% Mismatches: 0  
 Query Match: 78.7% Indels: 175  
 DB: 5 Gaps: 2

US-10-825-632-8 (1-1083) x ADR41398 (1-587)

QY 164 ATTATGTTACATCCCTATGTCGAAACAAGAGGCGCAGATTTCATTCGGTTATCCTAAA 223  
 DB 5 ILeHsValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLys 24  
 QY 224 ACAGGTACAGCAAAATCCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATTCATGCTGAA 283  
 DB 25 -----ThrAlaAsnProLysValThrPheLysMetSerGluLeuMetileAspAlaGlu 42  
 QY 284 CGAGGATCATA----- 295  
 DB 43 GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGlu 62  
 QY 295 ----- 295  
 DB 63 GlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIle 82  
 QY 295 ----- 295  
 DB 83 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 102  
 QY 295 ----- 295  
 DB 103 ProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 122  
 QY 295 ----- 295  
 DB 123 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 142  
 QY 295 ----- 295  
 DB 143 HisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLys 162  
 QY 295 ----- 295  
 DB 163 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 182  
 QY 295 ----- 295  
 DB 183 SerSerGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 202  
 QY 296 -----GTTGATCAA 304  
 DB 203 IleThrSerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 222  
 QY 305 GTCAAGAGCTGGTATATTTTGAAGGCACCAAGATCCCTTTTAGATCATCAGCTGAC 364  
 DB 223 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyr 242  
 QY 365 GTACTGATTCAGTAATCTCGAGGTCACAGGCTGACTGACCGTGGCTACTCAT 424  
 DB 243 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 262  
 QY 425 TCTTGTGTCATCAGTCAGCACTGTGATCTCTTTATAGTAAGTATAGTAACACAGAGAAT 484  
 DB 263 SerCysCysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsn 282  
 QY 485 CCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCTGTAAGATGACCCCAACTTGCAAAACA 544

Db 283 ProHisCysValSerLeuTyrLysLeuSerProGluAspAspProThrCysLysThr 302  
 QY 545 AAGGAATTTTGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCA 604  
 DB 303 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 322  
 QY 605 GAAATTTCTCTTTTGAAGTACTACTGGATTATTCATTGTATGGATGCTCTCAAGCCT 664  
 DB 323 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 342  
 QY 665 CATGATCTACAGCCTGGGAAGAAATATCCCTACTGCTGCTTCATATATATGCTGCTCAG 724  
 DB 343 HisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyrGlyGlyProGln 362  
 QY 725 GTGCAAGTTGGTCAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCC 784  
 DB 363 ValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAla 382  
 QY 785 TCTTAGTATTGTGGTTGTAGTCATAGACAACAGGGGATCTGTCCACCGAGGCTTAAA 844  
 DB 383 SerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLys 402  
 QY 845 TTTGAGGCGCTTTAAATATATAATGGTCAATAGAAATTCACGATCAGGTGGGAAGA 904  
 DB 403 PheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGly 422  
 QY 905 CTCCAATATCTAGCTTCTCGATATGATTTCATTGACTTAGATCGTGGGCATCCACGCG 964  
 DB 423 LeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGly 442  
 QY 965 TGGTCTTAGGAGGATACCTCTCCCTGATGGCATTATATGCAGAGGTGAGATATCTTCAGG 1024  
 DB 443 TrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArg 462  
 QY 1025 GTTGCTATATGTCGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGATACAG 1081  
 DB 463 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 481

RESULT 15  
 AAE14337  
 ID AAE14337 standard; protein; 580 AA.  
 AC AAE14337;  
 XX  
 DT 07-MAR-2002 (first entry)  
 XX  
 DE Human protease PRTS-2 protein.  
 XX  
 KW Human; protease; PRTS-2; tranquiliser; gene therapy; vaccine; allergy;  
 KW infection; dermatitis; arteriosclerosis; rheumatoid arthritis; hepatitis;  
 KW atherosclerosis; psoriasis; Alzheimer's disease; mental disorder; cancer;  
 KW gastrointestinal disorder; Cushing's syndrome; seizure; glaucoma; stroke;  
 KW epithelial disorder; urticaria; anorexia; trauma; asthma; eczema; nausea;  
 KW hypertension; neurological disorder; Parkinson's disease; drug screening;  
 KW cardiatic; cell proliferative disorder; multiple sclerosis; osteoporosis;  
 KW diabetes mellitus; glomerulonephritis; cardiovascular disorder; anaemia;  
 KW autoimmune disorder; inflammatory disorder; myocardial infarction; AIDS;  
 KW developmental disorder; reproductive disorder; infertility; diarrhoea;  
 KW dementia; acidosis; cataract; gynaecomastia; epilepsy; jaundice.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200183775-A2.  
 PD  
 PF 08-NOV-2001.  
 XX  
 PF 04-MAY-2001; 2001WO-US014651.  
 XX  
 PR 04-MAY-2000; 2000US-0202082P.  
 PR 11-MAY-2000; 2000US-0203566P.  
 PR 17-MAY-2000; 2000US-0205803P.  
 PR 25-MAY-2000; 2000US-0207477P.





```

Db      441 AsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrllysMet--- 459
QY      875 CAAATAGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTTC 934
Db      459 ----- 459
QY      935 ATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCCCTATGAGGATACCTCTCCCTGATG 994
Db      459 ----- 459
QY      995 GCATTAAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCCAGTCACTCTG 1054
Db      460 -----ValAlaIleAlaGlyAlaProValThrLeu 469
QY      1055 TGGATCTTCTATGATACAGGATACAG 1081
Db      470 TrpIlePheTyrlAspThrGlyTyrlThr 478

```

Search completed: May 2, 2006, 01:03:33  
Job time : 188.411 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 01:03:49 ; Search time 7.34981 Seconds  
(without alignments)  
2835.522 Million cell updates/sec

Title: US-10-825-632-8  
Perfect score: 1938  
Sequence: 1 ggaagaagatgccagatcg.....tatgatcacaggatacacgga 1083

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DRV=xlp  
-Q=/abs/ABSWEB.spool/US10825632/runat 01052006 105948 3262/app query.fasta\_1  
-DB=PIR -QFMT=fasta -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pcet -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p  
-USPR=US10825632.@CGN 1.1.92.@runat 01052006 105948 3262 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DSV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : PIR 80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	527.5	27.2	738	2	A87516 dipeptidyl peptida
2	491.5	25.4	931	2	T32919 hypothetical prote
3	430.5	22.2	741	2	JC5142 x-Pro dipeptidyl-p
4	369	19.0	711	2	S66261 dipeptidyl-peptida
5	315	16.3	766	1	CDHU26 dipeptidyl-aminope
6	312.5	16.1	793	2	T41703 dipeptidyl-peptida
7	309	15.9	760	1	S23752 dipeptidyl-peptida
8	291.5	15.0	792	1	A39914 dipeptidyl-aminope
9	280.5	14.5	818	1	A30107 dipeptidyl-aminope
10	250	12.9	931	2	A49737 dipeptidyl-peptida
11	245	12.6	795	2	T82858 hypothetical prote
12	240	12.4	829	2	T19514 dipeptidyl-aminope
13	226.5	11.7	683	2	E87495 dipeptidyl-aminope
14	221	11.4	803	2	A41793

15	220	11.4	709	2	B82580 alanyl dipeptidyl
16	219	11.3	779	2	T25173 hypothetical prote
17	219	11.3	799	2	T25174 hypothetical prote
18	205.5	10.6	631	2	H75007 probable acylamino
19	204	10.5	642	2	C71137 hypothetical prote
20	203	10.5	803	2	I68600 dipeptidyl aminope
21	203	10.5	865	2	IS4331 dipeptidyl aminope
22	199	10.3	622	2	F71174 hypothetical prote
23	197.5	10.2	743	2	T37700 probable dipeptidyl
24	196	10.1	657	2	E70025 probable acylamino
25	193	10.0	632	2	E75057 peptidase PAB1418
26	189	9.8	536	2	F90299 acylaminoacyl-pept
27	174	9.0	759	2	I38593 fibroblast activat
28	166.5	8.6	674	2	B84381 acylaminoacyl-pept
29	165	8.5	591	2	H74474 probable acylamino
30	162.5	8.4	569	2	S74053 probable acylamino
31	149	7.7	676	2	C97775 acylamino-acid-rel
32	147.5	7.6	667	2	A87711 prollyl oligopeptid
33	147.5	7.6	721	2	T09631 probable acylamino
34	140.5	7.2	591	2	A99444 acylaminoacyl-pept
35	140.5	7.2	606	2	T35378 probable peptide h
36	139.5	7.2	745	2	T33751 hypothetical prote
37	138.5	7.1	598	2	F84199 hypothetical prote
38	137.5	7.1	654	2	AD3183 peptidase [impute
39	134.5	6.9	732	1	JC4655 acylaminoacyl-pept
40	131.5	6.8	732	1	S07624 acylaminoacyl-pept
41	130.5	6.7	572	2	F72455 probable acylamino
42	129.5	6.7	555	2	AD1794 acylase and dieste
43	127.5	6.6	555	2	AB1419 hypothetical prote
44	126.5	6.5	623	2	T15945 prollyl oligopeptid
45	126.5	6.5	689	1	JN0585

ALIGNMENTS

RESULT 1  
A87516  
dipeptidyl peptidase IV [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: A87516  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: A87516  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-738 <STO>  
A:Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI000000C7616; GB:AE005673; NID:G13423647; P  
C:Genetics:  
A:Gene: CC2154

Alignment Scores:	1.08e-39	Length:	738
Pred. No.:	527.50	Matches:	137
Score:	41.2%	Conservative:	61
Percent Similarity:	28.5%	Mismatches:	132
Best Local Similarity:	27.2%	Indels:	151
Query Match:	2	Gaps:	13
DB:			
US-10-825-632-8 (1-1083) x A87516 (1-738)			
QY	5	GAAGATCCGATCAGCTGGCTGCTTGGTTCTCCAAAGAAATTCATAGATAT	64
	:::		
	189	LyspAlaLeuSerPheGlyValAlaGluPheIleValGlnGluLeuAspArgPhe	208
DB			
QY	65	TCTGCTATTGGTGTCTCCAAAAGCTGAAACAATCCCAAGTGGTGGTAAATTTTGA	124
DB	209	ThrGlyTyrTrpTrpSerProAspGluSer-----	219



```
QY 259 ----- 259
Db 400 SerAspGlyThrThrValTrpValGlnValMetSerArgAspGlnAlaGlnCysSerLeu 419
QY 259 ----- 259
Db 420 LeuLeuIleProTyrThrAspPheLeuLeuProGluGluLeuGlyGlySerIleLeuSglu 439
QY 259 ----- 259
Db 440 AspAsnLeuGlnLeuSerThrAspLeuAsnMetGlyValTrpAspAspLysSerHieGlu 459
QY 259 ----- 259
Db 460 GluThrMetGluLysProProArgGlyLysLeuArgGlyThrValGlnIleHisLysAla 479
QY 259 ----- 259
Db 480 ArgAsnAspTyrTrpIleAsnThrHisAsnAlaIleTyrProLeuLysIleThrAspGlu 499
QY 259 ----- 259
Db 500 GluHisProMetTyrGluPheIleTyrCysLeuGluLysProAsnGlySerCysLeuAla 519
QY 260 ----- TCAGAATAATGATT 274
Db 520 LeuIleSerAlaGluLeuAspGlnAsnGlyTyrCysArgHisThrGluGluLysLeuLeu 539
QY 275 GATGCTGAAGGA ----- AGGATCATAGTTGATGAGTACAGG 313
Db 540 MetAlaGluAsnPheSerIleAsnLysSerMetGlyIleValValAspGluValArgGlu 559
QY 314 CTGGTATATTTGAGGACCAAGACTCCCTTTAGACATCACCCTGTACGTAGTCAGT 373
Db 560 LeuValTyrTyrValAlaAsnGluSerHisProThrGluTrpAsnIle---CysValSer 578
QY 374 TACGTAAATCTCGAGAGGTGACAGCGTACTGACGCGTGGCTACTCACATTCTTGCTGC 433
Db 579 HisTyrArgThrGlyGlnHisAlaGlnLeuThrGluSerGlyIle----- 593
QY 434 ATCAGTCAGCACGTGCTACTTCTTTATAAGTAAGTATAGTAACCAAGG----- 481
Db 594 -----Cys-----PheLysSerGluArgAlaAsnGlyLysLeuAlaLeuAsp 607
QY 482 AATCCACACATGTGTGTCCTTTTACAAGCTATCA-----AGTCCTGAAGATGACCCAACT 535
Db 608 LeuAspHisGlyPheAlaCysTyrMetThrSerValGlySerProAlaGlu----- 624
QY 536 TGCAAAACAAGGAATTT---TGG----- 556
Db 625 CysArgPheTyrSerPheArgTrpLysGluAsnGluValLeuProSerThrValTyrAla 644
QY 557 GCCACCATTTTGGATTCAGCAGTCTCT---CTTCTGAC-----TATCTCTCCAGAA 607
Db 645 AlaAsnIleThrValSerGlyHisProGlyGlnProAspLeuHisPheAspSerProGlu 664
QY 608 ATTTTCTCTTTTGAAGT---ACTACTGGATTATACATTGTATGGGATGCTCTACAGCCT 664
Db 665 MetIleGluPheGlnSerLysLysThrGlyLeuMetHisTyrAlaMetIleLeuArgPro 684
QY 665 CATGATCTACAGCTCGAAGAAATATCTCTGCTGCTTTCATATATAGTGGTGGTCTCAG 724
Db 685 SerAsnPheAspProTyrLysLysTyrProValPheHisTyrValTyrGlyGlyProGly 704
QY 725 GTGCAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCC 784
Db 705 IleGlnIleValHisAsnAspPheSerTrpIleGlnTyrIleArg-----PheCys 721
QY 785 TCTCTAGTTATGTGGTTGTAGTATAGACACAGGGGATCTGTACCCAGGCTTAA 844
Db 722 ArgLeuGlyTyrValValPheIleAspAsnArgGlySerAlaHisArgGlyIleGlu 741
```

```
QY 845 TTTGAAGCGCCTTTAATATATAAATGGTCAAAATAGAAATGACATCAGGTGGAAGA 904
Db 742 PheGluArgHisIleHisLysLysMetGlyThrValGluValGluAspGlnValGluGly 761
QY 905 CTCCAATATCTAGCTTCTCGATAT---GATTTCAATTAGCTAGTGTGGCATCCAC 961
Db 762 LeuGlnMetLeuAlaGluArgThrGlyGlyPheMetAspMetSerArgValValHis 781
QY 962 GCCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATCTTC 1021
Db 782 GlyTrpSerTyrGlyTyrMetAlaLeuGlnMetIleAlaLysHisProAsnIleTyr 801
QY 1022 AGGGTGTCTATGCTGGGGCCCGCCAGTCACCTCTGTCGATCTCTATCATACAGGATACAG 1081
Db 802 ArgAlaIleAlaGlyAlaValSerAspTrpArgLeuTyrAspThrAlaTyrThr 821

RESULT 3
JC5142
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Xanthomonas maltophilia
C:Species: Xanthomonas maltophilia
C>Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5142
R:Kobashima, T.; Ito, K.; Yoshimoto, T.
J. Biochem. 120, 1111-1117, 1996
A:Title: Dipeptidyl peptidase IV from Xanthomonas maltophilia: Sequencing and expression
A:Reference number: JC5142; MUID:97164011; PMID:9010758
A:Accession: JC5142
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <KAB>
A:Cross-references: UNIPROT:P95782; UNIPARC:UPI0000085F8A; DDBJ:D83263; NID:g1753196; PI:
C:Comment: This enzyme catalyzes the hydrolytic removal of amino-terminal dipeptidyl res
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; membrane bound
F:4-18/Domain: transmembrane #status predicted <TM>
F:610/Active site: Ser #status predicted
F:685/Active site: Asp #status predicted
F:717/Active site: His #status predicted

Alignment Scores:
Pred. No.: 7,09e-31 Length: 741
Score: 430.50 Matches: 125
Percent Similarity: 38.2% Conservative: 57
Best Local Similarity: 26.2% Mismatches: 144
Query Match: 22.2% Indels: 151
DB: 2 Gaps: 13

US-10-825-632-8 (1-1083) x JC5142 (1-741)

QY 23 GGAGTCGCTACTTTGTTCTCCAGAGAATTTGTAGATATTTCTGGCTATTGGTGGTGT 82
Db 198 GlyIleAlaGluPheValAlaAspGluMetAspArgHisThrGlyTyrTrpAla 217
QY 83 CCAAAAGCTGAACAACTCCCGAGTGGTGGTAAATTTCTTAGAATTTCTATATCAAGAAAT 142
Db 218 ProAspSerAla-----IleAlaTyrAlaArgIle 228
QY 143 GATGAATCTGAGGTG-----GAAATATTCACTGTTACATCCCTATG 184
Db 229 AspGluSerProValProValGlnLysArgTyrGluValTyrAlaAspArgThrAspVal 248
QY 185 TTGGAACAACAGGAGGCGAGATTCCATTCCTTAAACAGGTCACAAATCCTAAA 244
Db 249 IleGluGln-----ArgTyrProAlaAlaGlyAspAlaAsnValGln 262
QY 245 GTCACTTTT----- 253
Db 263 ValLysLeuGlyValIleSerProAlaGluGlnAlaGlnThrClnTrpIleAspLeuGly 282
QY 254 AAGATGTCAGAAATA----- 268
Db 283 LysGluGlnAspIleTyrLeuAlaArgValAsnTrpArgAspProGlnHisLeuSerPhe 302
```









Db 305 SerLeuTrpHisAsnGluProLeuAlaGluProValValGlnAsnValLeuTrpValAsn 324  
QY 269 -----ATG 271  
Db 325 ThrSerSerValLeuValGlnPheThrAsnArgAsnSerThrCysIleThrAlaArgLeu 344  
QY 272 ATGGATCTCAAGGAAGGATCATTA 295  
Db 345 LeuAspThrGluLeuLysSerIleHisThrValLysThrGluCysLeuGluGluGlyTrp 364  
QY 295 ----- 295  
Db 365 TyrGluValGlnGlnSerAlaLysMetPheProLeuAsnAsnSerLeuValTrpGluAsn 384  
QY 296 -----GTTGATGAAGTCAGAAAGCTG- 316  
Db 385 TrpSerAspGlyTyrPheAspIleLeuAlaLeuAspAspTyrAsnHisLeuAlaPheIle 404  
QY 316 ----- 316  
Db 405 ProPheAsnGlySerSerProIleTyrLeuThrSerGlyAlaTrpAspValThrAspGly 424  
QY 317 -----GTATATTTTGAAGGACCAAGATCCCT 346  
Db 425 ProIleHisIleAspGlyAspPheGlyAsnValTyrPheLeuAlaThrLeuLysAspSer 444  
QY 347 TTAGAGCATCACCTGTAGTCAGTACGTAAATCCTGGAGAGGTGACAAAGCTGACT 406  
Db 445 ThrGluArgHisLeuTyrTyrValSer---LeuAspThrLeuGluIleTyrGlyIleThr 463  
QY 407 GACCGTGGCTACTCA-----CATTCCTGCTGCATCAGTCAGCAGCTGTGACTCTTT 457  
Db 464 AsnAsnGlyGluAspGluGlyTyrTyrSerThrSerPheSerProPheGlyAspPheTyr 483  
QY 458 ATAGTAAGTATAGTAAACAGAAAGATCCACACTGTGTGTCCTTTTCAAGCTATCAAGT 517  
Db 484 ValLeuAsnTyrHisGlyProAspValPro-----TrpGlnGluLeuArgSer 499  
QY 518 CCTGAAGATGACCCACTGCAAAACAAAGGAATTTGGGCCACCATTTGGATTACGCA 577  
Db 500 ThrLysAspLysAspTyrCysLeuSerLeuGluThrAsnSerArgLeuLysGlnGln--- 518  
QY 578 GGTCTCTCTGACTATACTCTCCAGAAAT-----TTCTCTTTTGAAGT 625  
Db 519 -----LeuSerSerIleThrLeuProSerValGluTyrGlyLysLeuThrPheAsnAsp 536  
QY 626 ACTACTGGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTACAGCCCTGGAAG 685  
Db 537 ThrThr---PheAsnPhe-----MetGluArgArgProArgAsnPheAspValAsnLys 553  
QY 686 AATATCTACTGTGCTGTTTCATATATATGTTGTCCTCAGCTGCTGCTGCTGAATAATCGG 745  
Db 554 LysTyrProValLeuPhePheAlaTyrGlyGlyProGlySerGlnGlnVal----- 570  
QY 746 TTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT----- 793  
Db 571 -----AlaLysLeuPheArgValAspPheGlnAlaTyrLeuAlaSerHisProAsp 587  
QY 794 -----TATGTGTTGTAGTATAGACAAACAGGGATCCTGTCAAGGAGCTTAAATTT 847  
Db 588 PheGluPheIleValThrLeuAspGlyArgGlyThrGlyPheAsnGlyAsnAlaPhe 607  
QY 848 GAAGCGCCTTTAAATATAAATGGGTCAAAATAGAAATGACATCAGCTGGAGGAGCTC 907  
Db 608 ArgTyrSerValSerArgHisLeuGlyGluTyrGluSerTyrAspGlnGlyGlnAlaGly 627  
QY 908 CAATATCTAGCTTCTCGATATGATTTTCATTGCTAGTCGTCGTCGTCATCCAGCTGG 967  
Db 628 LysPheTrpAlaAsp---LeuProPheValAspGluAsnHisValGlyIleTrpGlyTrp 646  
QY 968 TCCTTATGGAGATACCTCTCCCTGATGGCATTAAATGACAGGCTCAGATATCTTCAGGGTT 1027  
|||||

Db 647 SerTyrGlyGlyTyrLeuThrLeuLysThrLeu---GluThrGlnAspValPheSerTyr 665  
QY 1028 GCTATTGTGGGGCCCGAGTCACTCTGTGATCTTCTATGATACAGGATACACG 1081  
Db 666 GlyMetAlaValAlaProValThrAspTrpArgLeuTyrAspSerValTyrThr 683  
RESULT 7  
S23752  
diptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse  
NAlternate names: CD26 alpha subunit; THAM alpha subunit  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S23752; A46465; A56030  
R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.  
J. Biol. Chem. 267, 2200-2208, 1992  
A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di  
A:Reference number: S23752; MUID:92129288; PMID:1370813  
A:Accession: S23752  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-760 <MAR>  
A:Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384  
R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;  
J. Immunol. 147, 447-454, 1991  
A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (diptidyl peptidase  
A:Reference number: A46465; MUID:91302787; PMID:1712807  
A:Accession: A46465  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <VIV>  
A:Cross-references: UNIPARC:UPI0000172A2E  
A:Experimental source: M14.T thymoma cells, Swiss nu/nu  
A:Note: sequence extracted from NCBI backbone (NCBIP:42236)  
R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.  
Biochemistry 33, 15204-15214, 1994  
A:Title: Structure of the mouse diptidyl peptidase IV (CD26) gene.  
A:Reference number: A56030; MUID:9502780; PMID:7999781  
A:Accession: A56030  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 746-760 <BER>  
A:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620  
C:Genetics:  
A:Gene: CD26  
C:Superfamily: diptidyl-peptidase IV  
C:Keywords: diptidylpeptide hydrolase; glycoprotein; transmembrane protein  
F:213,223,315,514,679/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:624,702,734/Active site: Ser, Asp, His #status predicted  
Alignment Scores:  
Pred. No.: 7,9e-20 Length: 760  
Score: 309.00 Matches: 128  
Percent Similarity: 37.1% Conservative: 57  
Best Local Similarity: 25.7% Mismatches: 154  
Query Match: 15.9% Indels: 160  
DB: 1 Gaps: 23  
US-10-825-632-8 (1-1083) x S23752 (1-760)  
QY 2 GAAGAAGATGCCAGATCAGCTGGAGTGCCTTTGTTCTCAAGAGAA---TTTGAT 58  
Db 184 GluGluAsnValIleTyrAsnGlyIleThrAspTrpValTyrGluGluValPheGly 203  
QY 59 AGATATCTGGCTATTGGTGGTGTCCAA---CCTGAACAACT 100  
Db 204 AlaTyrSerAlaLeuTrpTrpSerProAsnAsnThrPheLeuAlaTyrAlaGlnPheAsn 223  
QY 101 CCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGAAATGATCAATCTCAGGTGGA 160  
Db 224 AspThrGlyValProLeuIleGluTyrSerPhe---TyrSerAspGluSer----- 239  
QY 161 ATTATTCATGTATCATCCCTATGTTGGAAACAAAGAGGCGACATTCATTCGTTATCCT 220  
|||||

Db 240 -----LeuGlnTyrProLysThrValTrpIleProTyrPro 251  
QY 221 AAAACAGGTACAGCAATCTCTAAAGTCATCTTT----- 253  
Db 252 LysAlaGlyAlaValAsnProThrValLysPhePheIleValAsnIleAspSerLeuSer 271  
QY 253 ----- 253  
Db 272 SerSerSerSerAlaIleProIleGlnIleProAlaProAlaSerValAlaAArgGlyAsp 291  
QY 254 ---AAGATGTCAGAAATATGATGATGCTGGAAGGAGGATCATAGTTGATGAAGTCAGA 310  
Db 292 HisTyrLeuCysAspValValTrpAlaThrGluGluArgIleSerLeuGlnTrpLeuArg 311  
QY 311 AGCTGCTATATTTGAAGGC-----ACCAAA 337  
Db 312 ArgIleGlnAsnTyrSerValMetAlaIleCysAspTyrAspLysIleAsnLeuTrp 331  
QY 338 GACTCCCTTTAGAG---CATCACCTGTACGTA-----GTGAGTTACGTA 379  
Db 332 AsnCysProSerGluGlnGlnHisValGluMetSerThrThrGlyTrpValGlyArgPhe 351  
QY 380 AATCCTGGAGGTGACAGGCTGACTGAC----- 409  
Db 352 ArgProAlaGluProHisPheThrSerAspGlySerSerPheTyrLysIleIleSerAsp 371  
QY 410 ---CGTGGCTACTCACATCTCTGCTGCATC-----AGTCAGCACTGTGACTTC--- 454  
Db 372 LysAspGlyTyrLysHisIleCysHisPheProLysAspLysLysAspCysThrPheIle 391  
QY 455 -----TTTATAGTAAG 466  
Db 392 ThrLysGlyAlaTrpGluValIleSerIleGluAlaLeuThrSerAspTyrLeuTyr 411  
QY 467 TATAGTAACAG-----AGATCCACACTGTGTCCTCCATTACAG----- 508  
Db 412 IleSerAsnGlnTyrLysGluMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuThr 431  
QY 509 -----CTATCAAGTCTCGAAGATGACCCCACTTGCAAA----- 541  
Db 432 AspHisThrAsnValLysCysLeuSerCysAspLeuAsnProGluArgCysGlnTyr 451  
QY 542 -----ACAAGGAAATTTGGGCCACCATTTGGATTGACAGAGTCCT---CTT 586  
Db 452 AlaValSerPheSerLysGluAlaLysTyrTyrGlnLeuGlyCysTrpGlyProGlyLeu 471  
QY 587 CQTGACTATACT----- 598  
Db 472 ProLeuTyrThrLeuHisArgSerThrAspHisLysGluLeuArgValLeuGluAspAsn 491  
QY 599 -----CCTCCAGAAATTTCTCTTT--- 619  
Db 492 SerAlaLeuAspArgMetLeuGlnAspValGlnMetProSerLysLysLeuAspPheIle 511  
QY 620 ---GAAAGTACTCTGATTTACATTTGATGGATGCTCTACAAGCCCTCATGATCTACAG 676  
Db 512 ValLeuAsnGluThrArgPhe---TrpTyrGlnMetIleLeuProProHis---PheAsp 529  
QY 677 CTTGGAAGAAATATCTACTGTGCTGTGTTTCATATATGTTGGTGTCTCAGGTGCGATTTGGTG 736  
Db 530 LysSerLysLysTyrProLeuLeuLeuAspValTyrAlaGlyProCysSerGln----- 547  
QY 737 AATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAT-----ACCCTAGCCTCT 787  
Db 548 -----LysAlaAspAlaSerPheArgLeuAsnTrpAlaThrTyrLeuAlaSer 563  
QY 788 CTAGGTTATGTGTTGTAGT---ATAGACACACAGGGATCTCTCACCGAGGCTTAAA 844  
Db 564 ThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLys 583  
QY 845 TTTGAGGGCCCTTTAAATATAAATGGGTCAAAATGAGAAATTTGACGATCAGGTGGAGGA 904  
Db 584 IleMetHisAlaIleAsnArgLeuGluGlyThrLeuGluValGluAspGlnIleGluAla 603

QY 905 CTCCAATATCTAGCTTCTCGATATGATTTCAATTGACTTATAGATCGTGTGGCATCCACCGC 964  
Db 604 AlaArgGlnPheVal---LysMetGlyPheValAspSerLysArgValAlaIleTrpGly 622  
QY 965 TGTCTCTATGAGGATCACTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAGG 1024  
Db 623 TrpSerTyrGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLys 642  
QY 1025 GTTCTATTTGCTGGGGCCCGCAGTCACTCTGTGGATCTTCTATGATACAGATACAG 1081  
Db 643 CysGlyIleAlaValAlaProValSerArgTrpGluTyrTyrAspSerValTyrThr 661  
RESULT 8  
A39914  
dipeptidyl-peptidase IV (EC 3.4.14.5), membrane-bound form precursor - rat  
N/Alternate names: GP110; membrane glycoprotein 110K; OX-61  
N/Contains: dipeptidyl-peptidase IV, soluble form  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004  
C/Accession: A39914; A33315; B33315; A60730; A42203; S38949; A31781  
R/Hong, W.; Doyle, D.  
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987  
A/Title: cDNA cloning for a bile canalicular domain-specific membrane glycoprotein of ra  
A/Reference number: A39914; MUID:88068516; PMID:3479775  
A/Accession: A39914  
A/Molecule type: mRNA  
A/Residues: 1-792 <HON>  
A/Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN:  
J/Ogata, S.; Misumi, Y.; Ikehara, Y.  
J. Biol. Chem. 264, 3596-3601, 1989  
A/Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA an  
A/Reference number: A33315; MUID:89123496; PMID:2563382  
A/Accession: A33315  
A/Molecule type: mRNA  
A/Residues: 1-37, A'39-182, I', 184-331, T', 333-351, C', 353-393, V', 395-561, L', 563-623,  
A/Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:G203973; PIDN:AAA41036.1; PID:  
R/McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.  
Hepatology 11, 534-544, 1990  
A/Title: Identification of the bile canalicular cell surface molecule GP110 as the ectop  
quence.  
A/Reference number: A60730; MUID:90228896; PMID:1970322  
A/Accession: A60730  
A/Molecule type: protein  
A/Residues: 28-47, XX', 50-53, 55-58 <MCC>  
A/Cross-references: UNIPARC:UPI0000172A38  
J/Ogata, S.; Misumi, Y.; Teuji, E.; Takami, N.; Oda, K.; Ikehara, Y.  
Biochemistry 31, 2582-2587, 1992  
A/Title: Identification of the active site residues in dipeptidyl peptidase IV by affini  
A/Reference number: A42203; MUID:92190188; PMID:1347701  
A/Accession: A42203  
A/Molecule type: protein  
A/Residues: 'R', 625-630, 'X', 632-648 <OG2>  
A/Cross-references: UNIPARC:UPI0000172A39  
R/Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.  
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993  
A/Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl p  
A/Reference number: S38949; MUID:94128239; PMID:7905271  
A/Accession: S38949  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 281-302 <IWA>  
A/Cross-references: UNIPARC:UPI0000172A3A  
R/Hong, W.; Doyle, D.  
J. Biol. Chem. 263, 16992-16998, 1988  
A/Title: Membrane orientation of rat gp110 as studied by in vitro translation.  
A/Reference number: A31781; MUID:89034185; PMID:3182821

Db	404	GluValIleSerIleGluAlaLeuThrSøAspTyrLeuTyrTyrIleSerAsnGluTyr	423
QY	479	---AAGAAATCCACACTGTGTGTCCTCTTTACAAG---	508
Db	424	LysGluMetProGlyGlyArgAsnLeuTyrIleGlnLeuThrAspHisThrAsnLys	443
QY	509	-----CTATCAAGTCCTGAAGTACGCCCACTTGCAAA-----ACA	544
Db	444	LysCysLeuSerCysAspLeuAsnProGluAygCysGlnTyrTyrSerValSerLeuSer	463
QY	545	AAGAAATTTGGGCCACCATTTTGGATTGACAGAGTCTCT---CTTCCTGACTATACT---	598
Db	464	LysGluAlaIstYrTyrGlnLeuGlyCysArgGlyProGlyLeuProLeuTyrThrLeu	483
QY	599	-----CCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA---	640
			503

[illegible]

Db QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY

543 LeuLeulleAspValTyAlaGlyProCysSerGln-----Lysala 556  
755 GTCAGATATTCCGGTGAAT-----ACCTTAGCCTCTCTAGGTATTGTGGTTGTA 805  
557 AspAlaIaPheArgPheAsnTrpAlaThrTyrlleuAlaSerThrgluAsnlieIleVal 576  
806 GTG---ATAGACAACAGGGGACTCTGTCCACGAGGGCTTAAATTTCAAGCGCGCTTTAAA 862  
577 AlaSerPheAspGlyArgGlySerGlyTyrlleuGlnGlyAspLysileMethisAlaIleAsn 596  
863 TATAAATGGGTCAAATAGAAATTGACGATCAGGTGGAGGAGCTC---CAATATCTTAGCT 919  
597 LysArgLeuGlyThrLeuGluValgluaspGlnIleGluAlaalargGlnPheLeu---- 615  
920 TCTCGATATGATTTTCATTGACTTAGATCTGTGTGGGATCCACGGCTGTGCCTATGGAGA 979  
616 ---LysMetGlyPheValaspSerLysGlnValAlaIlelrpGlylrpSerTyrlleuGly 634  
980 TACCTCTCCCTGATGGCATTAATGACAGAGTGACAGATATCTTCAGGGTTCGTTCTGGGG 103  
635 TyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaVal 654  
1040 GCCCAGTCACCTCTGTGATCTTCTATGATACAGATACACG 1081  
655 AlaIrpValSerArgTrpGluTyrlleuAspSerValTyrlleuThr 668

RESULT 9  
A30107  
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (*Saccharomyces cerevisiae*)  
N;Alternate names: protein YHR028c  
C;Species: *Saccharomyces cerevisiae*  
C;Date: 07-Jun-1990 #sequence\_revision 30-May-1997 #text\_change 09-Jul-2004  
C;Accession: S46780; A30107  
R;Du, Z.  
submitted to the EMBL Data Library, June 1994  
A;Description: The sequence of *S. cerevisiae* cosmid 8082.  
A;Reference number: S46773  
A;Accession: S46780.  
A;Molecule type: DNA  
A;Residues: 1-818 <DUZ>  
A;Cross-references: UNIPROT:P18962; UNIPARC:UPI00000031ASf; EMBL:U10399; NID:95  
R;Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.  
J. Cell Biol. 108, 1363-1373, 1989  
A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase

**THIS PAGE BLANK (USPTO)**







A;Reference number: S45451; MUID:95066382; PMID:7975897		Db	
A;Accession: S45451		533 LysGluProIleGlnLeuThrLysGlyAsnTrpGluValThrGlyAsnGlyValGly 552	
A;Molecule type: DNA		QY	
A;Residues: 1-931 <ANN>		296 GTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGAGACTCCCTTTAGAGCAT 355	
A;Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:g347196; PIDN:AAA351119.1; R:Gallison, F.; Dujon, B.		Db	
submitted to the EMBL Data Library, October 1995		553 TyrGluTyrGluThrAspThrIlePhePheThrAlaAsnGluIleGlyValMetSerGln 572	
A;Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome		QY	
A;Reference number: S60938		356 CACCTGTACGTAGTCAGTTAGTA----- 379	
A;Accession: S60946		Db	
A;Molecule type: DNA		573 HisLeuTyrSerIleSerLeuThrAspSerThrThrGlnAsnThrPheGlnSerLeuGln 592	
A;Residues: 1-931 <GAL>		QY	
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; R:Boyer, J.; Fairhead, C.; Gaillon, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon, B.		Db	
submitted to the Protein Sequence Database, July 1996		380 AATCTCTGGAGAGTGACCAAGGCTCAGCCGTGGCTACTTCACATTTCTTGTGTCATCACT 439	
A;Reference number: S67104		QY	
A;Accession: S67112		593 AsnPro-----SerAspLysTyrAspPheTyrAspPheGluLeuSer 606	
A;Molecule type: DNA		Db	
A;Residues: 1-931 <BOY>		440 CAGCACTGTGACTCTCTTTTATAAGTATAGTAACACGAGGAATCCA----- 487	
A;Cross-references: UNIPARC:UPI0000136060; EMBL:Z75127; NID:g1420507; PIDN:CAA99437.1; R:Gallison, F.; Dujon, B.		QY	
Experimental source: strain S288C		607 SerSerAlaArgTyrAlaIleSerLysLysLeuGlyProAspThrProIleLysValAla 626	
Yeast 12, 877-885, 1996		Db	
A;Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV c		QY	
A;Reference number: S71713; MUID:96437977; PMID:8840505		488 -----CACTGTGTGCTCCCTTTTACAAGCTA 511	
A;Accession: S71721		Db	
A;Status: nucleic acid sequence not shown; translation not shown		627 GlyProLeuThrArgValLeuAsnValAlaGluIleHisAspSerIleLeuGlnLeu 646	
A;Molecule type: DNA		QY	
A;Residues: 1-931 <GAW>		512 TCAAGTCTCTGAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCACCACTTTTGGAT 571	
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1; R:Gallison, F.; Dujon, B.		Db	
Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995		QY	
C;Genetics:		632 -----GGATTTTACATTTATGGGATGCTCTACAGGCTCTATGATCTACAGCCCTGGA 682	
A;Gene: SGD:STE13; YC11		QY	
A;Cross-references: SGD:S0005745; MIPS:YOR219c		671 LeuAspAspGlyValGluIleAsnTyrIleGluIleLysProAlaAsnLeuAsnProLys 690	
A;Map position: 15R		Db	
C;Function:		QY	
A;Description: involved in processing of alpha-factor prepropheromone		683 AGAAATATCTCTACTGTCTGTTCATATATGCTGCTCTCAGGTGCGTGGTGAATAAT 742	
F;113-150/Domain: transmembrane #status predicted <TM>		QY	
F;377/Binding site: carbohydrate (Asn) (covalent) #status predicted		691 LysLysTyrProIleLeuValAsnIleTyrGlyProGlySerGlnThrPheThrThr 710	
F;785,863,896/Active site: Ser, Asp, His #status predicted		QY	
Alignment Scores:		743 CGGTTTAAAGGAGTCAGATTTTCGCTTGAATACCTCCTAGCTCTCTAGGTGTATGTGGTT 802	
Pred. No.:	1.87e-14	Length:	931
Score:	250.00	Matches:	101
Percent Similarity:	38.3%	Conservative:	58
Best Local Similarity:	24.3%	Mismatches:	162
Query Match:	12.9%	Indels:	94
DB:	2	Gaps:	15
US-10-825-632-8 (1-1083) x A49737 (1-931)		Db	
QY		44 CAGAGATTTGATAGATATCTCGCTATTGGTGTCTCA-----AAAGCTGAA 94	
Db		434 GlnLysAspSerIleLeuTyrAsnGlyLysTrpIleSerProAspThrPheArgPheGlu 453	
QY		95 ACAATCCCGAGTGGTGAATTTCTTAGAATTTCTATAGAGAAATGATGATCTGAG 154	
Db		454 IleThrAspArgAsnSerLysIleLeuAspValLysValTyrAspIleProSerSerGln 473	
QY		155 GTGGAAATATTATCATGTTACATCC----- 178	
Db		474 MetLeuThrValArgAsnThrAsnSerAsnLeuPheAsnGlyTrpIleGluLysThrLys 493	
QY		179 -----CCTATGTTGGAACACAGAGG----- 199	
Db		494 AspileLeuSerIleProProLysProGluLeuLysArgMetAspTyrGlyTyrIleAsp 513	
QY		200 -----GCAGATTCA-----TTCCGTTATCTCTAAACACAGGTACAGCA 235	
Db		514 IleHisAlaAspSerArgGlyPheSerHisLeuPheTyrTyrPro-----ThrValPheAla 532	
QY		236 AATCTCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGCTGAAGGAAGATCAT 295	

RESULT 11  
F82858  
dipeptidyl-peptidase XF0015 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: F82858  
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717; PMID:10910347

A>Note: for a complete list of authors see reference number A59328 below

A/Accession: F82958  
A/Status: Preliminary  
A/Molecule type: DNA  
A/Residues: 1-795 <SIM>  
A/Cross-references: UNIPROT:Q9PH9; UNIPARC:UPI00000C2287; GB:AB003856; GB:AB003849; NID  
B/Simpon, A.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facinani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, R.C.; Miyaki, C.Y.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A/Reference number: A59328  
A/Contents: annotation  
C/Genetics:  
A/Gene: XF0015

Alignment Scores:  
Pred. No.: 5,23e-14 Length: 795  
Score: 245.00 Matches: 98  
Percent Similarity: 41.2% Conservative: 49  
Best Local Similarity: 27.5% Mismatches: 140  
Query Match: 12.6% Indels: 70  
DB: 2 Gaps: 16

US-10-825-632-8 (1-1083) x F82858 (1-795)

```
QY 62 TATTTCGCTATTGCTGTCGCAAAAGCTGAAACACTCCAGTGGTGAATTCCT 121
Db 396 PheAsnApphEgLyTrpLeuProspAsnArgThr-----Leu 408
QY 122 AGAATCTCTATGAGAAATGATCTGAGGTGAATATTCATGTTACATCCCT 181
Db 409 TrpLeuSerGluGluSerGlyTrpSerHisLeuThrTrpLeu----- 424
QY 182 ATGTTGGAACAGGAGGCGCATTCATCCGTTATCTTAAACAGGTACCAATCCT 241
Db 425 -----GluThrGlyTrpProlyGln 431
QY 242 AAAGTCATTTT-----AAGATCTCAGAAATAATGATGCTGAAGGAGGATC 292
Db 432 Arg-----ThrPheGlyHisTrpGluValSerAlaProArgPheSerAlaAspGly----- 448
QY 293 ATAGTTGATGAAGTCAGAAAGCTGTGTATATTTTGAAGGCCAACCAAGACTCCCTTTAGAG 352
Db 449 -----ArgSerValTrpPhe----- 453
QY 353 CATCACCTGTACGTAGTCAGTTAGTAAATCCTGGA-----GAGTGACAAAGGTG--- 403
Db 454 -----ValCysAsnProThrTrpProGlyGlyTrpGluValCysLysLeuAsp 469
QY 404 ACTGACCGTGCTACTCACATCTTGTGTCATCAGTCAGTCACTGT---GACTTCTTTATA 460
Db 470 LeuAspSerGlyHisPheAspThrLeuThraLeuGlnGlyValGluAspPheAspLeu 489
QY 461 AGTAAGTATAGTACCAAGAGATCCACACTGTGTCTCCCTTTTACAGCTATCA---AGT 517
Db 490 SerProaspGlyArgGln-----IleLeuValLeuTrpSerGlySerTrpLeu 505
QY 518 CCTGAAGATGACCACTGCAAAAACAAAGGAATTTGGGCCACCATTTTG----- 568
Db 506 ProGlnLeuAlaValIleSerThrAenGlyGlyGlnAlaArgValLeuThrAspThr 525
QY 569 ---GATTTCAGCAGTCTCTCTCTGACTATCTCTCCAGGAATTTCTCTTTGAAAGT 625
Db 526 ArgThrSerThrPheLysAlaLysGlnTrpIleGlnProGlyTrpValGlnIleProSer 545
```

```
QY 626 ACTACTGGATTT---ACATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCTTGA 682
Db 546 LysHisGlyAlaGlyValIleTrpGlyLysTrpGlyProAlaSerLeuGlnProGlu 565
QY 683 AAGAAATATCTACTGCTGCTGTATATATGCTGCTCTCAGGTGCTGCTGCTGAATAAT 742
Db 566 LysArgTrpProIleValMetPheValHisGlyAlaGlyTrpLeuGlnAsnValSerAla 585
QY 743 CGGTTTAAAGGATCAAGTATTTCCGCTTGAATACCTAGCTCTCTA----- 790
Db 586 ArgTrpPro-----SerTrpPheArgGluGlnMetPheHisThrLeuValGlnLys 603
QY 791 GGTATGCTGCTAGTACTAGTACACACAGGGGATCTCTCCAGGAGCTTAAATTTGAA 850
Db 604 GlyTrpIleValLeuAspLeuAspTrpArgAlaSerAlaGlyHisGlyArgAspTrpArg 623
QY 851 GGGCCTTTAAATATAAATGCTCAATAGAAATGATGATCAGTGGGAGGACTCCAA 910
Db 624 ThrAlaIleTrpArgAsnMetGlyHisProGluLeuGluAspTrpLeuAspGlyLeuAsp 643
QY 911 TATCTAGCTTCTCGATATGATTTCATTCAGTATAGTCTGTGGCATCCAGCGTGGTCC 970
Db 644 TrpLeuValAlaHisLysGlnGly---AspArgArgAlaGlyMetTrpGlyGlySer 662
QY 971 TATGAGGATACCTCTCCCTGATGATGATTAATGACAGAGGTTCAGATATCTTCAGGGTGTCT 1030
Db 663 TyrGlyGlyPheMetThrTrpMetAlaMetPheArgAlaProGlyThrPheLysAlaGly 682
QY 1031 ATTGCTGGGCGCCAGTCACTCTGTGATCTTCTATGATATACAGGATACAGC 1081
Db 683 AlaAlaLeuArgProValGlyAspTrpMetGlnTrpAsnHisGluTrpThr 699
```

RESULT 12  
T19514  
hypothetical protein C27C12.7 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T19514  
R/Thomas, K.  
submitted to the EMBL Data Library, March 1996  
A/Reference number: Z19134  
A/Accession: T19514  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-829 <WIL>  
A/Cross-references: UNIPROT:Q18253; UNIPARC:UPI000007DC53; EMBL:Z69883; PIDN:CAA93743.1;  
A/Experimental source: clone C27C12  
C/Genetics:  
A/Gene: CESP:C27C12.7  
A/Map position: X  
A/Introns: 4/3; 51/1; 123/3; 156/3; 188/2; 279/2; 392/3; 501/3; 553/1; 583/1; 606/2; 647/2;  
C/Superfamily: dipeptidyl-peptidase IV

Alignment Scores:  
Pred. No.: 1.5e-13 Length: 829  
Score: 240.00 Matches: 107  
Percent Similarity: 36.4% Conservative: 77  
Best Local Similarity: 21.1% Mismatches: 144  
Query Match: 12.4% Indels: 178  
DB: 2 Gaps: 20

US-10-825-632-8 (1-1083) x T19514 (1-829)

```
QY 23 GGAGTCGCTACCTTTGTTGTTCTCCCAAGAGAA---TTTGATAGATATTTCTGGCTATTGGTG 79
Db 249 GlyLeuAlaAsnTrpLeuTrpGluGluIleLeuGluAlaSerSerAlaValTrpTrp 268
QY 80 TGTCAAAGCTGAAACAACTCCCACTGGTGGTGAATTTCTTAGAATTTCTATATGAAGAA 139
Db 269 -----SerProSerGlyArgTrpValSerTrpLeuArgPheAspAsp 282
QY 140 AATGATGAATCTGAGGTGGAAATTTATTTCATGCCCTTATGTTACATCCCTTATGTTGAAACAAAGGAG 199
```

[illegible]

Db 626 Gin1leValSerGlnTyrAspIleAlaTyr-----AlaArgIleAspValArgGlyThr 643  
 QY 827 TGTCAACCGAGCGCTTAAATTTGAAGCGCGCTTTAAATATATAAAATGGGTCAATAGAAATT 886  
 Db 644 GlyGlyArgGlyTyrAspValValGluAlaValTyrArgLeuGlyAspAlaGluVal 663  
 QY 887 GACGATCAGGTGGAGGACCATCAATATCTAGCTTCTCGATATGATTCATTGACTTATAGT 946  
 Db 664 ValAspThrLeuAspMetIleArgAlaPheIleAsnThrPheGlyPheIleAspGluAsp 683  
 QY 947 CGTGTGGGATCCAGCGCTGCTCTATGAGGATACCTC---TCCCTGATGGCATTAATG 1003  
 Db 684 ArgIleAlaValMetGlyTyrSerTyrGlyGlyPheLeuThrSerLysIleAlaLeuLys 703  
 QY 1004 CAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCCAGTCACCTCTGTGGATCTTC 1063  
 Db 704 AspGlnGlyGluLeuValLysCysAlaIleSerIleAlaProValThrAspPheLysTyr 723  
 QY 1064 TATGATACAGGATACACG 1081  
 Db 724 TyrAspSerAlaTyrThr 729

RESULT 13  
 E87495  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C:Accession: E87495  
 R:Nierman, W.C.; DeBooy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D. D.; B.; Laub, M.T.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heide-  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fra-  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
 A:Reference number: AB7249; MUID:21173696; PMID:11259647  
 A:Accession: E87495  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-683 <STO>  
 A:Cross-references: UNIPROT:Q9AGU7; UNIPARC:UPI00000C7585; GB:AE005673; NID:913  
 C:Genetics:  
 A:Gene: CC1986

Alignment Scores:		
Pred. No.:	2,47e-12	Length: 683
Score:	226.50	Matches: 66
Percent Similarity:	45.9%	Conservative: 28
Best Local Similarity:	32.2%	Mismatches: 94
Query Match:	11.7%	Indels: 17
DB:	2	Gaps: 6

US-10-825-632-8 (1-1083) x E87495 (1-683)  
 QY 377 GTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCCTGCTGCATC 436  
 Db 356 ValLysThrGlyLysValSerAlaLeuThrGlyGluGlyHisValThrAlaPheAspVal 375  
 QY 437 AGTCAGCACTGTGACTCTCTTTATAGTAAGTATAGTAACACAGAGAATCCACACTGTGTG 496  
 Db 376 GlyProSerGlyIleValPheAlaSer-----AspSerLeuLysSerProSer----- 391  
 QY 497 TCCCTTTTACAAGCTATCAAGTCCTCGAGATGACCCCAACTTGGCAAAACAAAGGAATTTTGG 556  
 Db 392 GluLeuPheLeuLeu-----ProAlaLysGlyProAlaValLysVal----- 405  
 QY 557 GCACACATTTTGGATTCAGCAGGTCCTCTCTGACTATATACTCTCTCCAGAAATTTTCTCT 616  
 Db 406 AlaSerValSerGluAlaLeuLysAspValAlaTrpGlyGluProGluGlnPheSer 425  
 QY 617 TTGGAAGTACTACTGGATTACATTGATCGGATGCTCTACAGGCTCATGATCTACAG 676  
 Db 426 PheLysGlyTyrAsnAspGluThrValHisGlyPheLeuValLysProAlaAsnPheAsp 445  
 QY 677 CTGGGAAGAAATATCTTACTGTGCTGTTTCATATATATGTTGCTCTCAGGTGCAGTTGGTG 736

```

Db      446  ProAlaLysLysTyPProValAlaPheLeuLleHisGlyGlyProGln----- 461
Qy      737  AATAATCGGTTTAAAGGNGTCAAGTATTTCGCGTTGAAT-----ACCTAGCTCTCTA 790
Db      462  --GlySerPheSerAsnAlaTrpSerTyAArgPAsnProGlnValTyAlaAsnAla 480
Qy      791  GGTATCTGTTGTAGTATAGACACACAGGATCTGTCACCGAGGCGCTTAAATTTGAA 850
Db      481  GlyTyAlaValMetIleAspPheHisGlySerThrGlyTyGlyGlnAlaPheThr 500
Qy      851  GGCCGCTTTAAATATAAATGAGTCAATAGAAATGACATCAGGTGGAGGACTCCAA 910
Db      501  AspSerIleSerArgHisTrpGlyAspArgProLeuGluAAspLeuGlnLysGlyTrpSer 520
Qy      911  TATCTAGCTTCTCGATATGATTTCATTGACTTGTAGATCTGTGGCATCCACGCGTGTCC 970
Db      521  PheValLeuSerLysTyGlyPheLeuAspGlyAspAlaCysAlaLeuGlyAlaSer 540
Qy      971  TATGGAGGATACCTC 985
Db      541  TyrGlyGlyTyrMet 545

RESULT 14
A:1793
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A41793
R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthhold, R.J.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
A:Title: Differential expression of two distinct forms of mRNA encoding members of a dip
A:Reference number: A41793; PMID:92108018; PMID:1729689
A:Accession: A41793
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-803 <WAD>
A:Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A83C; GB:M76429; NID:g408719; PIDN:
A:Note: sequence extracted from NCBI backbone (NCBIP:751138)
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein
F:257,342/Binding site: carbohydrate (asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 7,95e-12 Length: 803
Score: 221.00 Matches: 103
Percent Similarity: 32.9% Conservative: 58
Best Local Similarity: 21.0% Mismatches: 169
Query Match: 11.4% Indels: 160
DB: 2 Gaps: 16

US-10-825-632-8 (1-1083) x A41793 (1-803)
Qy      23  GGAGTCGCTACCTTTGTTCTCAAGAGAA-----TTTGATAGATATTCTGGCTATTGGTGG 79
Db      225  GlyLeuSerAspTrpLeuTyGlyGluGluLeuLysThrHisIleAlaHisTrp 244
Qy      80  TGTCCTCAAAAGCTGAACAACTCCAGTGGTGAATTTCTAGAAATCTATATGAAGAA 139
Db      245  SerProAspGly-----ThrArgLeuAlaTyAlaThr 255
Qy      140  AATGATGATCTGAGTGGAAATTTTCATGTT-----ACATCCCTCATGTTGGAA 190
Db      256  IleAsnAspSerArgValProValMetGluLeuProThrTyThrGlySerValTyPro 275
Qy      191  ACAAGGAGGCGAGATTTCATTCGTTATCTTAAACAGGTCACAGCAATCTTAAATCTACT 250
Db      276  Thr-----AlaLysProTyHisTyProLysAlaGlyCysGluAsnProSerIleSer 293
Qy      251  TTT----- 253
Db      294  LeuHisValIleGlyLeuAsnGlyProThrHisAspLeuGluMetThrProProAspAsp 313

254 ---AAGATGTCAGAAAATAATGATT----- 274
314 ProArgMetArgGluTyTyTrileThrMetValLysTrpAlaThrSerThrLysValAla 333
275 -----GATGCT----- 280
334 ValAsnTrpLeuSerArgAlaGlnAsnValSerIleLeuThrLeuCysAspAlaThrThr 353
280 ----- 280
354 GlyValCysThrLysLysHisGluAspGluSerGluAlaTrpLeuHisArgGlnAsnGlu 373
281 -----GAAGAAGG----- 289
374 GluProValPheSerLysAspGlyArgLysPhePhePheValArgAlaIleProGlnGly 393
289 ----- 289
394 GlyGlnGlyLysPheTyHisIleThrValSerSerSerGlnProAsnSerSerAsnAsp 413
290 -----ATCATAGTTGATGAA 304
414 AsnIleGlnSerIleThrSerGlyAspTrpAspValThrLysIleLeuSerTyAspGlu 433
305 GTCAGAGAGCTGGTATATTTCAGAGCACCAGCACTCCCTTTAGAGCATCACCCTGTAC 364
434 LysArgSerGlnIleTyPheLeuSerThrGluAspLeuProArgArgGlnLeuTy 453
365 GTAGTCAGTTTACGTAATCCTCGAGAGGTGACAGG-----CTGACT 406
454 SerAlaSerThrVal-----GlySerPheAsnArgGlnCysLeuSerCysAspLeuVal 471
407 GACCGTGGCTACTCACATCTTCTGTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAG 466
472 AspAsnCysThrTyPheSerAlaSerPheSerProGlyAlaAspPhePheLeuLeuLys 491
467 TATAGTAACAGAGAATCCACACTGTGTCTCCCTTTACAGCACTACAGTCTCTGAGAT 526
492 CysGluGlyProGlyValPro---ThrValSerValHisAsnThrThrAspLysLysLys 510
527 GACCCCAACTTGCAAAACAAAGGAA---TTTTGGGCCACCATTTTGGATTGACGAGGTCT 593
511 MetPheAspLeuGluThrAsnGluHisValGlnLysAlaIleSerAspArgGlnMetPro 530
584 CTTCTCTGACTATACCTCTCCAGAAATTTCTCTTTTGAAGAGTACTACTGGATTTCATG 643
531 LysValGluTyArgLysIleGlu-----ThrAspAspTyAsnLeu 544
644 TATGGGATGCTCTACAGCCTCATGATCTACAGCCTGCAAGAAATATCTACTGTGCTG 703
545 ProIleGlnIleLeuLysProAlaThrPheThrAspThrAlaHisTyProLeuLeuLeu 564
704 TTCATATATGTTGGTGGTCTCAGGTGCTGTAATATCGGTTTAAAGAGTCAAGTAT 763
565 ValValAspGlyThrProGlySerGlnSerValAlaGluLysPhe-----AlaValThr 582
764 TTCGCTTGAATACCTAGCTCTCTAGGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 823
583 TrpGluThrValMetValSerSerHisGlyAlaValValValLysCysAspGlyArgGly 602
824 TCCTGTCCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAATAATATGAGTCAAG 883
603 SerGlyPheGlnGlyThrArgLeuLeuHisGluValArgArgArgLeuGlySerLeuGlu 622
884 ATTGACGATCAGGTGGAGGAGCACTCCAATATCTAGCTTCTCGATATGATTTCATTTGAT 943
623 GluLysAspGlnMetGluAlaValArgValMetLeu---LysGluProTyTrileAspLys 641
944 GATCTGTGGGATCCACGCTGCTCTATGGAGGATPACTCTCTCCCTGATGATGATGATG 1003
642 ThrArgValAlaValPheGlyLysAspTyGlyTyLeuSerThrTyTrileLeuLeuPro 661
1004 CAGAGGTGAGAT-----ATCTTCAGGTTGCTATTGCTGGGCGCCCAAGTCACT 1051
```

Db 662 AlaLysGlyAspGlyGlnAlaProValPheSerCysGlySerAlaLeuSerProIleThr 681  
QY 1052 CTGTGATCTTCTATCATACAGATACAGC 1081  
Db 682 AspPheLysLeuTyAlaSerAlaPheSer 691

RESULT 15  
B82580  
alanyl dipeptidyl peptidase XP2260 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: B82580  
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717; PMID:10910347  
A;Note: for a complete list of authors see reference number A59328 below  
A;Accession: B82580  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-709 <SIM>  
A;Cross-references: UNIPROT:Q9PB84; UNIPARC:UPI00000C29B1; GB:AE004038; GB:AE003849; NID  
A;Experimental source: strain 9a5c  
R;Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A;Reference number: A59328  
A;Contents: annotation  
C;Genetics:  
A;Gene: XP2260

Alignment Scores:  
Pred. No.: 9.66e-12 Length: 709  
Score: 220.00 Matches: 73  
Percent Similarity: 41.0% Conservative: 32  
Best Local Similarity: 28.5% Mismatches: 99  
Query Match: 11.4% Indels: 52  
DB: 2 Gaps: 9

US-10-825-632-8 (1-1083) x B82580 (1-709)

QY 263 GAATATGATGATGCTGAGGAGGATCATAGTTGATGAAGTCAGAGCGCTGGTATAT 322  
Db 358 GluIleAlaLeuSerAlaAspGlyLysAlaLeu-----Tyr 369  
QY 323 TTTCAGGACCAAGACTCCCTTTAGAGACTACCTGTAGTGTAGTCAGTTACGTAAAT 382  
Db 370 ValAsnAla-----AspAspHisGlyGluHisProLeuPheLysValAsp---IleAla 386  
QY 383 CTGTGAGAGGTGACAGGCTGACTGACCGGTGCTACTCATCTTCTGTGCTCAGTCAG 442  
Db 387 SerGlyLysValGluLysTrpValGlyGluGlySerValHisAlaProValLeuAlaGly 406  
QY 443 CACTGTGACTCTTTATAGTAAGTATAGTAACACGAGAGATCCACACTGTGTCTCCTT 502  
Db 407 -----GlyLysLeuAlaPheAlaArgasn----- 414  
QY 503 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTTGGGCCACC 562  
Db 415 ---SerLeuLysSerAlaAsp-----GlnIlePheValThr 425  
QY 563 ATTTGGATTCAGCAGGCTCTCTCTGTGACTATACCTCTCCA----- 604

Search completed: May 2, 2006, 01:38:29  
Job time : 58.7491 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 00:42:23 ; Search time 49.8855 Seconds  
(without alignments)  
3063.366 Million cell updates/sec

Title: us-10-825-632-8  
Perfect score: 1938  
Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatcacgga 1083

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues  
Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abss/ABSSWEB.spool/US10825632/runat\_01052006\_105946\_3226/app\_query.fasta\_1  
-DB=UniProt -QFMT=faetan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-OUTFMT=ptc -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-DUALIGN=200 -THR SORT=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p  
-USER=US10825632\_@CGN\_1\_1\_694\_@runat\_01052006\_105946\_3226 -NCPU=6 -ICPU=3  
-NO MAP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt\_05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1836.5	94.8	898	1 DPP8_HUMAN	O6v1x1 homo sapien
2	1756.5	90.6	892	1 DPP8_MOUSE	Q80ya7 mus musculu
3	1220.5	63.0	863	1 DPP9_HUMAN	Q86t12 homo sapien
4	1204.5	62.2	862	1 DPP9_MOUSE	Q8bvg4 mus musculu
5	1183.5	61.1	923	2 Q4SBM6_TETNG	Q4sbm6 tetraodon n
6	1182.5	61.0	847	2 Q6GR22_XENLA	Q6gr22 xenopus lae
7	944	48.7	508	2 Q75273_HUMAN	Q75273 homo sapien
8	718.5	37.1	1053	2 Q9VC20_DROME	Q9vc20 drosophila
9	718.5	37.1	1113	2 Q9VC19_DROME	Q9vc19 drosophila
10	705	36.4	886	2 Q7QBK1_ANOGA	Q7gbk1 anopheles g
11	684	35.3	740	2 Q5TTK8_ANOGA	Q5ttk8 anopheles g
12	630	32.5	621	2 Q7PTT8_ANOGA	Q7ptt8 anopheles g
13	605.5	31.2	803	2 Q54U01_DICDI	Q54u01 dictyosteli
14	601.5	31.0	557	2 Q5TXJ2_ANOGA	Q5txj2 anopheles g
15	527.5	27.2	738	2 Q9A6E0_CAUCR	Q9a6e0 caulobacter
16	494	25.5	763	2 Q8EAB7_SHEON	Q8eab7 shewanella

17	491.5	25.4	927	2	Q955K3_CABEL	Q955k3 caenorhabdi
18	491.5	25.4	931	2	O44987_CABEL	O44987 caenorhabdi
19	474	24.5	745	2	O6F317_9P8ED	O6f317 pseudomonas
20	471	24.3	751	2	O4UPD3_XANCP	O4upd3 xanthomonas
21	471	24.3	751	2	Q8P3V8_XANCP	Q8p3v8 xanthomonas
22	468	24.1	745	2	O5H5W8_XANOR	O5h5w8 xanthomonas
23	465.5	24.0	895	2	O61CU7_CABER	O61cu7 caenorhabdi
24	463	23.9	743	2	O5QX36_IDILO	O5qx36 idiomarina
25	463	23.9	757	2	O8PFD7_XANAC	O8pfd7 xanthomonas
26	454	23.4	749	2	Q7NEK8_GLOVI	Q7nek8 gloeobacter
27	448	23.1	766	2	O4TNP1_9SPHN	O4tnp1 erythrobact
28	440	22.7	596	2	Q6K880_ORYSA	Q6k880 oryza sativ
29	430.5	22.2	741	2	P95782_XANNA	P95782 xanthomonas
30	422.5	21.8	746	2	O9FNF6_ARATH	O9fnf6 arabidopsis
31	418.5	21.6	778	2	O51ZP7_MAGGR	O51zp7 magnaporthe
32	409.5	21.1	739	2	O5FOY6_GLUOX	O5fgy6 gluconobact
33	403.5	20.8	720	2	O5NNM8_ZYMON	O5nnm8 zymomonas m
34	403.5	20.8	732	2	Q7MUM6_PORGI	Q7muw6 porphyromon
35	403	20.8	736	2	Q8A028_BACTN	Q8a028 bacteroides
36	390.5	20.1	552	2	Q8GUJ7_ARATH	Q8guj7 arabidopsis
37	383	19.8	719	2	O5LJ01_BACFN	O5lj01 bacteroides
38	382	19.7	719	2	Q65J00_BACFR	Q65j00 bacteroides
39	372	19.2	736	2	O5LGU5_BACFN	O5lgu5 bacteroides
40	372	19.2	736	2	Q64XP9_BACFR	Q64xp9 bacteroides
41	370.5	19.1	732	2	Q8A2Q1_BACTN	Q8a2q1 bacteroides
42	369	19.0	711	2	Q47900_FLAME	Q47900 flavobacter
43	366.5	18.9	730	2	Q93JY4_9BACT	Q93jy4 prevotella
44	365	18.8	809	2	Q4NVV1_9DBLT	Q4nvv1 anaeromyxob
45	353.5	18.2	731	2	Q6L872_PREIN	Q6l872 prevotella

ALIGNMENTS

RESULT 1  
DPP8\_HUMAN STANDARD; PRT; 898 AA.  
ID DPP8\_HUMAN STANDARD; Q724D3; Q724B1; Q81WG7; Q8NEM5; Q96JX1; Q9HEM2;  
AC Q6V1X1; Q7Z4C8; Q724D3; Q724B1; Q81WG7; Q8NEM5; Q96JX1; Q9HEM2;  
AC Q9HEM3; Q9HEM4; Q9HEM5; Q9NXP4;  
DT 13-SEP-2005 (Rel. 48, Created)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII) (DP8)  
DE (Prolyl dipeptidase DPP8) (Dipeptidyl peptidase IV-related protein 1)  
DE (DPP8-1).  
GN Name=DPP8; Synonyms=DPP8P1; ORFNames=MSTP097, MSTP135, MSTP141;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF  
RP 334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM  
RP 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION  
RP CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION,  
RP AND SUBCELLULAR LOCATION.  
RC TISSUE=Placenta;  
RX MEDLINE=20467194; PubMed=11012666;  
RA Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W.,  
RA Correll M.D.;  
RT "Cloning, expression and chromosomal localization of a novel human  
RT dipeptidyl peptidase (DPP) IV homolog, DPP8.";  
RL Eur. J. Biochem. 267:6140-6150(2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME  
RP REGULATION, BIOPHYSICO-CHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND  
RP TISSUE SPECIFICITY.  
RC TISSUE=Testis;  
RX PubMed=12662155; DOI=10.1042/BJ20021914;  
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akinaanya K.O.;  
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member  
RT of an emerging subgroup of serine proteases.";  
RL Biochem. J. 373:179-189(2003).





*Mus musculus* (Mouse)

480	Tyrl	yslle	Thr	Ser	lle	Leu	ys	Gl	u	Ser	Lys	Tyr	Lys	Arg	Ser	Ser	Gly	Gly	Leu	Pro	490	
295	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	295	
500	Ala	Pro	Ser	Asp	Phe	lys	Cys	Pro	lle	ys	Gl	u	lle	Ala	lle	Thr	Ser	Gly	Gl	u	Trp	519
296	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	322	
520	Gl	u	Val	Leu	Gly	Arg	His	Gly	Ser	Asn	lle	Gln	Val	Asp	Gl	u	Val	Arg	Arg	Leu	Val	539
323	TTT	GAG	GCC	ACC	AA	GAC	TCCC	TTT	AG	AGC	ATC	AC	CTT	CT	AG	CT	AG	T	C	AG	T	382
540	Phe	Gl	u	Gly	Thr	lys	Asp	Ser	Pro	Leu	Gl	u	His	His	Leu	Tyr	Val	Val	Ser	Tyr	Val	559
383	CCT	CGA	GAG	TGA	CA	AG	CGT	CA	CTG	CGG	TCT	AC	ATT	TCT	TG	CT	G	C	AT	C	AG	442
560	Pro	Gly	Gl	u	Val	Thr	Arg	Leu	Thr	Asp	Arg	Ser	Gly	Tyr	Ser	His	Ser	Cys	Cys	lle	Ser	579
443	CAC	TGT	GACT	TCT	TTT	TAA	GTA	AGT	AT	AGT	AT	AGT	AT	AGT	AT	AGT	AT	AGT	AT	AGT	AT	502
580	His	Cys	Asp	Phe	lle	Ser	Lys	Tyr	Ser	Asn	Gln	Lys	Asn	Pro	His	Cys	Val	Ser	Leu	Leu	599	
503	TAC	AA	GCT	ATC	CA	AGT	CTC	TGA	AG	TAC	ACC	CA	CTT	G	CA	AA	CA	CA	AG	AA	TTT	562
600	Tyr	lys	Leu	Ser	Ser	Pro	Gl	u	Asp	Asp	Pro	Thr	Cys	Lys	Thr	Lys	Gl	u	Phe	Trp	Ala	619
563	ATT	TTT	GGA	TT	CAG	CAG	TCT	CTC	TCT	CTG	ACT	AT	TAC	TCT	CC	CA	GAA	AA	TTT	TCT	TTT	622
620	lle	Leu	asp	Ser	Ala	Gly	Pro	Leu	Pro	Asp	Tyr	Thr	Pro	Pro	Gl	u	lle	Phe	Ser	Phe	Gl	639
623	AGT	ACT	ACT	GGA	TTT	ACA	TTG	TAT	GGA	TG	CTCT	CA	CA	AG	CTC	AT	GAT	CTC	CA	AG	CTC	682
640	Ser	Thr	Thr	Gly	Phe	Thr	Leu	Tyr	Gl	u	Met	Leu	Tyr	Lys	Pro	His	Asp	Leu	Gln	Pro	Gly	659
683	AAG	AA	AT	ATC	CT	CA	GTG	TG	TCT	CA	TAT	AT	G	GG	TG	CTC	CAG	GTG	CAG	TTG	TG	742
660	Lys	lys	tyr	Pro	Thr	Val	Leu	Phe	lle	tyr	Gly	Pro	Gln	Val	Gln	Leu	Val	Asn	Asn	Asn	679	
743	CGG	TTT	AA	GAG	GAT	CA	AGT	AT	TTT	CCG	CT	CA	AT	AC	CTC	AG	CTC	CTC	AG	CTT	AT	802
680	Arg	Phe	lys	Gly	Val	lys	Tyr	Phe	Arg	Leu	Asn	Thr	Leu	Ala	Ser	Leu	Gly	Tyr	Val	Val	699	
803	GTA	GTG	TAT	GAT	AG	CA	CA	G	GG	AT	CCT	GTG	C	AC	G	GG	GCT	TAA	TTT	CA	AG	862
700	Val	Val	lle	Asp	Asn	Arg	Gly	Ser	Cys	His	Arg	Gly	Leu	lys	Phe	Gl	u	Arg	Ala	Phe	lys	719
863	TAT	AAA	TGG	CGT	CA	AA	TAG	AAA	TTG	AC	CA	AT	CAG	CT	GGA	GCA	CTC	CA	AA	TAT	CT	922
720	Tyr	lys	Met	Gly	Gln	lle	Gl	u	lle	Asp	Asp	Gln	Val	Gl	u	Gl	u	Leu	Gln	Tyr	Leu	739
923	CGA	TAT	GAT	TTT	CAT	TG	ACT	TAG	AT	CGT	G	GG	CA	TCC	ACG	CGT	G	GT	CTC	TAT	G	982
740	Arg	Tyr	Asp	Phe	lle	Asp	Leu	Asp	Arg	Val	Gly	lle	His	Gly	Trp	Ser	Tyr	Gly	Gly	Tyr	759	
983	CTC	TCC	CTG	TGG	CA	TAA	TTC	AG	CA	GCT												

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=C57BL/6J; Tissue=Testis; DOI=10.1038/nature01266;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kakuwaka T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nozaki R., Schombach C., Gojohori T.,  
 RA Baldarelli R., Hilli D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schrim L.M., Kanagiri A., Matsuda H., Batalov S., Beisels K.W.,  
 RA Blake J.A., Brad T., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gliss C., Godzik A., Gough J.,  
 RA Grimmond S., Guertlich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Megloun D.R., Maitais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verrado R., Wagner L., Walsh E.D., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyszynski B., Yanagisawa M., Yang L., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Haashizume W., Imotani K., Iehi Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=C57BL/6; Tissue=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Straube R.L., Feingold E.A., Grouse L.H., Derse J.G., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Casavant T.B., Scheetz T.E.,  
 RA Raha S., Locantore N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Besak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [3]  
 RP PHOSPHORYLATION SITE TYR-325.  
 RX PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;  
 RA Shu H., Chen S., Bi Q., Mumbly M., Brekken D.L.;  
 RT "Identification of phosphoproteins and their phosphorylation sites in  
 RT the WEHI-231 B lymphoma cell line";  
 RL Mol. Cell. Proteomics 3:279-286 (2004).  
 CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal  
 CC dipeptides from proteins having a pro or ala residue at position  
 CC 2. May play a role in T-cell activation and immune function (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-

Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided  
 Zaa is neither Pro nor hydroxyproline.  
 CC -!- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine  
 CC proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride  
 CC (AEBSF), and by dl-isopropylfluorophosphate. Specifically inhibited  
 CC by isocycline derivatives (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 DR EMBL; AK016546; BAB30295.2; -; mRNA.  
 DR EMBL; BC043124; AAH43124.1; -; mRNA.  
 DR EMBL; BC059222; AAH59222.1; -; mRNA.  
 DR MEROPS; S09.018; -.  
 DR Ensembl; ENSMUSG0000032393; Mus musculus.  
 DR MGI; MGI:1921638; Dpp8.  
 DR InterPro; IPR001375; Peptidase\_S9.  
 DR InterPro; IPR002469; Peptidase\_S9B.  
 DR InterPro; IPR000379; Ser\_estrs.  
 DR Pfam; PF00930; DPPIV\_N; 1.  
 DR Pfam; PF00930; Peptidase\_S9; 1.  
 DR AminoPeptidase; Hydrolase; Phosphorylation; Protease; Serine protease.  
 FT ACT\_SITE 749 749 Charge relay system (By similarity).  
 FT ACT\_SITE 827 827 Charge relay system (By similarity).  
 FT ACT\_SITE 859 859 Charge relay system (By similarity).  
 FT MOD\_RES 325 325 Phosphotyrosine.  
 FT CONFLICT 87 87 G -> R (in Ref. 1).  
 SQ SEQUENCE 892 AA; 102186 MW; 59081CD9792E03ED CRC64;  
 Alignment Scores:  
 Pred No.: 2,840-147 Length: 892  
 Score: 1756.50 Matches: 343  
 Percent Similarity: 65.7% Conservative: 7  
 Percent Local Similarity: 64.4% Mismatches: 10  
 Query Match: 90.6% Indels: 173  
 DB: 1 Gaps: 1  
 US-10-825-632-8 (1-1083) x DPP8\_MOUSE (1-892)  
 QY 2 GAAGAAGATCCAGATCAGTCGATCGCTACCTTTGTTCTCAAGAAGAAATTTGATAGA 61  
 Db 254 GUGUGAASPProArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 273  
 QY 62 TATCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCACGTCGTTGTTAAATTCCT 121  
 Db 274 TyrSerGlyTyrTrpTrpCysProGlnAlaGluArgThrProSerGlyGlyLysLeu 293  
 QY 122 AGAATCTATATCAAGAAATGATGAATCTGAGGTGGAATATTATTCATGTTACATCCCT 181  
 Db 294 ArgileLeuTyrGluGluAsnAspGluSerGluValGluilelleHieValThrSerPro 313  
 QY 182 ATCTTGAAACAGAGGAGGAGCAGATTCATTCGTTATCTTAAACAGGTACAGCAATCCT 241  
 Db 314 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 333  
 QY 242 AAAGTCACCTTTTAAAGATGTCAGAAATAATGATTGATGTCGTAAGGAGGATCAT- 295  
 Db 334 LysValThrPheLysMetSerGluilelleValAlaAspAlaGlyGlyLysLeuAspVal 353  
 QY 295 ----- 295  
 Db 354 IleAspLysGluLeuValGlnProPheGluLeuLeuPheGluGlyValGluTyrIleAla 373  
 QY 295 ----- 295  
 Db 374 ArgAlaGlyTyrThrProGluGlyLysHieAlaTrpSerIleLeuLeuAspArgSerGln 393  
 QY 295 ----- 295

Db	394	ThrHisLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspAla	413	Db	754	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	773
Qy	295	-----	295	Qy	1043	CCAGTCACTCTGGGATCTTCTATGATACAGATACAGC	1081
Db	414	MetAspArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr	433	Db	774	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr	786
Qy	295	-----	295	RESULT 3			
Db	434	GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnThr	453	ID	DP99 HUMAN	STANDARD; PRT; 863 AA.	
Qy	295	-----	295	AC	Q8GTI2; O6A137; Q6UAL0; Q6ZMT2; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WDXD;		
Db	434	GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnThr	453	AC	Q9NT8; Q9BVR3;		
Qy	295	-----	295	DT	13-SEP-2005 (Rel. 48, Created)		
Db	454	HisGluAspGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu	473	DT	13-SEP-2005 (Rel. 48, Last sequence update)		
Qy	295	-----	295	DE	Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DP9)		
Db	474	TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro	493	DE	(Dipeptidyl peptidase-like protein 9) (DPLP9) (Dipeptidyl peptidase		
Qy	295	-----	295	GN	IV-related protein 2) (DRP-2).		
Db	494	AlaProSerAspPheLysCysProIleLysGluGluIleThrIleThrSerGlyGluTrp	513	GN	Name=DP9; Synonyms=DRP2;		
Qy	295	-----	295	OS	Homo sapiens (Human).		
Db	494	AlaProSerAspPheLysCysProIleLysGluGluIleThrIleThrSerGlyGluTrp	513	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Qy	295	-----	295	OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
Db	494	AlaProSerAspPheLysCysProIleLysGluGluIleThrIleThrSerGlyGluTrp	513	OC	Homo.		
Qy	296	-----GTTGATGAAGTCAGAAGCTGGTATAT	322	OC	NCBI_TaxID=9606;		
Db	514	GluValLeuGlyArgHisGlySerAsnIleTrpValAspGluAlaArgLysLeuValTyr	533	RN	[1]		
Qy	323	TTTGAAGGCACCAAGACTCCCTTACAGCATACCTGACGTAGTCAGTCAGTAAAT	382	RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.		
Db	534	PheGluGlyThrLysAspSerProLeuGluHisIleLysLeuTyrValThrSerTyrAlaAsn	553	RX	MEDLINE=22347328; PubMed=12459266; DOI=10.1016/S0378-1119(02)01059-4;		
Qy	383	CTGGAGAGGTGACAGGCTGACTGACCGTGGTACTCACATCTTGTCTGATCAGTCAG	442	RA	Olsen C., Wagtmann N.;		
Db	554	ProGlyGluValValArgLeuThrAspArgGlyTyrSerHisSerCysLeuSerArg	573	RT	"Identification and characterization of human DPP9, a novel homologue		
Qy	443	CACGTGACATCTTTTATAAGTAACTAGTAACTACACAGAAATCCACACTGTGTGTCCTT	502	RT	of dipeptidyl peptidase IV.";		
Db	574	HisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu	593	RL	Gene 299:185-193(2002).		
Qy	503	TAAAGTATCAAGTCTGAAGATGACCCCACTTCCAAACAAAGAAATTTGGCCACC	562	RN	[2]		
Db	594	TyrLysLeuSerProGluAspAspProValHisLysThrLysGluPheTrpAlaThr	613	RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), CATALYTIC ACTIVITY,		
Qy	563	ATTTCGATTCACAGCTCCTCTCTGACTATCTCTCCAGAAATTTCTCTTTTGA	622	RP	BIOPHYSICOCHEMICAL PROPERTIES, ENZYME REGULATION, TISSUE SPECIFICITY,		
Db	614	IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu	633	RP	AND SUBCELLULAR LOCATION.		
Qy	623	AGTACTACTGGATTTACATTGTATGGATGCTCTACAAGCCTCATGATCTACAGCTGGA	682	RC	TISSUE=Colon;		
Db	634	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	653	RX	PubMed=12662155; DOI=10.1042/BJ20021914;		
Qy	683	AAGAAATATCTACTGTGCTTCATATATGTTGCTCTCAGGTGCTAGTGGTAATAT	742	RA	Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akineanya K.O.;		
Db	654	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsn	673	RA	"Cloning and characterization of dipeptidyl peptidase 10, a new member		
Qy	743	CGTTTAAAGAGTCAAGTATTCGCTTGATACCTAGCTCTCTAGGTGTTAGTGTT	802	RT	of an emerging subgroup of serine proteases.";		
Db	674	ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal	693	RL	Biochem. J. 373:179-189(2003).		
Qy	803	GTAGTGATACACACAGGGATCTCTGACCCAGGGCTTAAATTTGAAGGCGCTTTAAA	862	RN	[3]		
Db	694	ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	713	RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).		
Qy	863	TATAAATGGGTCAAAATGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT	922	RC	TISSUE=Placenta, and Skin;		
Db	714	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	733	RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
Qy	923	CGATATGATTCATTTAGCTAGTCTGTGGGCATCCAGCTGCTCTATGGAGGATAC	982	RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
Db	734	GlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr	753	RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
Qy	983	CTCTCCCTGATGCATTAATGACAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCC	1042	RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
		-----		RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,		
		-----		RA	Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,		
		-----		RA	Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,		
		-----		RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
		-----		RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
		-----		RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
		-----		RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
		-----		RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
		-----		RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
		-----		RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
		-----		RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
		-----		RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
		-----		RA	Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,		
		-----		RA	Schmerch A., Schein J.E., Jones S.J.-M., Marra M.A.;		
		-----		RT	"Generation and initial analysis of more than 15,000 full-length human		
		-----		RT	and mouse cDNA sequences.";		
		-----		RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		



Db	254	ThrGlyTyrTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu	723	Db	614	HisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnPro	633
Qy	122	AGAAATTCATATGAGAAATGATGAATCTGAGGTGGAATATTATTCATGTTACATCCCT	181	Qy	680	CGAAGAAATATCTACTGTCCTTCATATATATGTTGGTCTCTCAGGTGAGTTGGTGAAT	739
Db	274	ArgIleLeuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro	293	Db	634	GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn	653
Qy	182	ATGTTGAAACAAAGAGGGCGAGTTCATTCGTTATCCCTAAACAGGTACACCAATCCT	241	Qy	740	AATCGGTTTAAAGAGTCAAGTATTTCCTTGAATACCTAGCCTCTCTAGGTTATGTG	799
Db	294	AlaLeuGluGluValArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro	313	Db	654	AsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla	673
Qy	242	AAAGTCACCTTTTAAAGATCTGACAAATAATGATTGATCTGAAGGAAGATCATA	295	Qy	800	GTTGTAGTATGATACACAGGGGATCTCTGTCACCGAGGCTTAATTTTGAAGCGCCTTT	859
Db	314	LysIleAlaLeuLysLeuAlaGluPheClnThrAspSerGlnGlyLysIleValSerThr	333	Db	674	ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu	693
Qy	295	-----	295	Qy	860	AAATATAAAATGGGTCAAAATAGAAATGACATCAGGTGGAAGGACTCCAATATCTAGCT	919
Db	334	GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla	353	Db	694	LysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAla	713
Qy	295	-----	295	Qy	920	TCTCGATATGATTTTCATTGACTTATGATCGTGTGGGCATCCACGGCTGGTCTATGGAGGA	979
Db	354	ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln	373	Db	714	GluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGly	733
Qy	295	-----	295	Qy	980	TACCTCTCCCTCATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGG	1039
Db	374	GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu	393	Db	734	PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGly	753
Qy	295	-----	295	Qy	1040	GCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACAGC	1081
Db	394	GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValTyr	413	Db	754	AlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr	767
Qy	295	-----	295	RESULT 4			
Db	414	GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer	433	DPP9 MOUSE			
Qy	295	-----	295	ID	DPP9 MOUSE	STANDARD;	PRT; 862 AA.
Db	434	GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis	453	AC	Q8BVG4: Q6KAM9; Q8BWT9;		
Qy	295	-----	295	DT	13-SEP-2005 (Rel. 48, Created)		
Db	454	LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer	473	DT	13-SEP-2005 (Rel. 48, Last sequence update)		
Qy	295	-----	295	DT	13-SEP-2005 (Rel. 48, Last annotation update)		
Db	474	ProGlyGluAspGluPheLysCysProIleLysGluGluLeuAlaLeuThrSerGlyGlu	493	DE	Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)		
Qy	295	-----	295	GN	Name=Dpp9;		
Db	494	TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuVal	513	OS	Mus musculus (Mouse)		
Qy	320	TATTTGAAGGCACCAAGACCTCCCTTTAGAGCATCAGTGTAGTACGTACGTACGTA	379	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Db	514	TyrPheGluGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGlu	533	OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
Qy	380	AATCTCGAGGTGACAGGCTGACTGACCGTGTACTCACATCTTCTGTCGATCAGT	439	OC	Muroidea; Muridae; Mus;		
Db	534	AlaAlaGlyGluIleValArgLeuThrProGlyPheSerHisSerCysSerMetSer	553	OX	NCBI_TaxID=10090;		
Qy	440	CAGCATGTGACTTCTTTTATAAGTATAGTATAACAGAGAATCCACACTGTGTGTCC	499	RN	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).		
Db	554	GlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrProProCysValHis	573	RC	STRAIN=C57BL/6J; TISSUE=Liver; and Olfactory bulb;		
Qy	500	CTTTACAGCTATCAAGTCTGAGTATGATGATGATGATGATGATGATGATGATGATGAT	559	RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;		
Db	574	ValTyrLysLeuSerGlyProAspAspProLeuHisLysLysLysLysLysLysLysLys	593	RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,		
Qy	560	ACCATTTTGATTCAGCAGGTCTCTTCTCTGACTACTACTCTCTCCAGAAATTTCTTTT	619	RA	Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,		
Db	594	SerMetMetGluAlaAlaSerCysProProAspTyrValProProGluIlePheHisPhe	613	RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,		
Qy	620	GAAGTACTACTGGATTTTACATTGATGGGATGCTCTTCAAGCCCTCATGATCTTACACCT	679	RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,		
		---		RA	Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,		
		---		RA	Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,		
		---		RA	Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,		
		---		RA	Gasteland T., Gariboldi M., Gissi C., Godzik A., Gough J.,		
		---		RA	Grimmond S., Gustinch S., Hirokawa N., Jackson I.J., Jarvis E.D.,		
		---		RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,		
		---		RA	Konagaya A., Kurochkin I.V., Lee Y., Lienhard B., Lyons P.A.,		
		---		RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,		
		---		RA	Nagashima T., Numa K., Okido T., Pavan W.J., Perte G., Pesole G.,		
		---		RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,		
		---		RA	Ravadin A., Schneider C., Sempé C.A., Setou M., Shimada K.,		
		---		RA	Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,		
		---		RA	Verardo L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,		
		---		RA	Wilmom R., Zavelan M., Zhu Y., Zimmer A., Carrincci P., Hayatsu N.,		
		---		RA	Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,		
		---		RA	Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,		
		---		RA	Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,		
		---		RA	Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,		







QY 295 ----- 295  
 DB 433 GluGlyGlnGlnAspPheCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 452  
 QY 295 ----- 295  
 DB 453 LeuTyrArgValThrValGluLeuLysThrLysAspTyrAspTyrThrGluProLeuSer 472  
 QY 295 ----- 295  
 DB 473 ProThrGluAspGluPheLysCysProLleLysGluGluValAlaLeuThrSerGlyGlu 492  
 QY 296 ----- 319  
 DB 493 TrpGluValLeuSerArgHisGlySerLysLleTyrValAsnGluGlnThrLysLeuVal 512  
 QY 320 TATTGAGGCCACCAAGACTCCCTTTAGAGCATCACCCTGCTAGTGTAGTACAGTACGTA 379  
 DB 513 TyrPheGlnGlyThrLysAspThrProLeuGluHisLysLeuTyrValValSerTyrGlu 532  
 QY 380 AATCCTGAGAGGTGACACAGGCTCACTGACCGCTACTCATCTTCTGTCATCAGT 439  
 DB 533 SerAlaGlyGluLleValArgLeuThrThrLeuGlyPheSerHisSerCysSerMetSer 552  
 QY 440 CAGCAGCTGTGATCTTTTATAAGTATAGTATAGTAACCAAGAAATCCACACTGTGTGCC 499  
 DB 553 GlnSerPheAspMetPheValSerHisTyrSerSerValSerThrProCysValHis 572  
 QY 500 CTTTACAGCTACTCAAGTCTGAGATGACCCAACTTGCAAAACAAAGAAATTTGGGCC 559  
 DB 573 ValTyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAla 592  
 QY 560 ACCATTGGAATTCAGCAGGTCTCTTCTGATATACCTACTACTCTCCAGAAATTTCTCTTT 619  
 DB 593 SerMetMetGluAlaAlaAsnCysProAspTyrValProGluLlePheHisPhe 612  
 QY 620 GAAATCTACTGAGTTTACATTTATGATGATGATGATGATGATGATGATGATGATGAT 679  
 DB 613 HisThrArgAlaAspValGlnLeuTyrGlyMetLleTyrLysProHisThrLeuGlnPro 632  
 QY 680 GGAAGAATAATCTTACTGCTGCTGCTTATATATGATGATGATGATGATGATGATGAT 739  
 DB 633 GlyArgLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn 652  
 QY 740 AATCGGTTPAAGAGTCAAGTATTTCCGTTGATACCTAGCTCTCTAGTTATG 799  
 DB 653 AsnSerPheLysGlyLleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla 672  
 QY 800 GTTGTAGTATAGACACAGGGGATCCTGTCACCGAGGCTTAAATTTGAAGCGCCTTT 859  
 DB 673 ValValAlaAspGlyArgGlySerCysGlnArgGlyLeuHisPheGluGlyAlaLeu 692  
 QY 860 AATATATAAATGGTCAATAGAAATGACATGAGTGGAGGAGCTCCATATCTAGCT 919  
 DB 693 LysAsnGlnMetGlyGlnValGluLeuGluAspGlnValGluGlyLeuGlnTyrValAla 712  
 QY 920 TCTCGATATGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 979  
 DB 713 GluLysTyrGlyPheLleAspLeuSerArgValAlaLleHisGlyTyrSerTyrGlyGly 732  
 QY 980 TACTCTCCCTGATGGCATTATATGAGAGGTGATATCTTTCAGGGTGTCTATGCTGGG 1039  
 DB 733 PheLeuSerLeuMetGlyLeuLleHisLysProGlnValPheLysValAlaLleAlaGly 752  
 QY 1040 GCCCAGTCTACTGTGGATCTTCTATGATACAGATACAG 1081  
 DB 753 AlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 766  
 RESULT 5  
 Q4SBM6\_TETNG  
 ID Q4SBM6\_TETNG PRELIMINARY; PRT; 923 AA.  
 AC Q4SBM6;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Chromosome 15 SCAP14667, whole genome shotgun sequence.  
 DE (Fragment).  
 GN ORFNames=GSTENG0020903001;  
 OS Tetraodon nigroviridis (green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 OC NCBI\_TaxID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
 RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
 RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,  
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J., Bosak S.,  
 RA Parra G., Lardier J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,  
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 RT the early vertebrate proto-karyotype."  
 RL Nature 431:946-957(2004).  
 [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Genoscope; Whitehead Institute Centre for Genome Research;  
 RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CA901014667; CAG01956.1; -; Genomic\_DNA.  
 FT NON\_TER 1  
 FT NON\_TER 923  
 FT SQUENCE 923 AA; 105211 MW; 139702382004009 CRC64;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2,88e-96 Length: 923  
 Score: 1183.50 Matches: 237  
 Percent Similarity: 49.1% Conservative: 59  
 Best Local Similarity: 39.3% Mismatches: 64  
 Query Match: 61.1% Indels: 243  
 DB: 2 Gaps: 2  
 US-10-825-632-8 (1-1083) x Q4SBM6\_TETNG (1-923)  
 QY 2 GAAGAAGATCCAGATCAGCTGGAGTCTACCTTTGTTCTCCAAAGAAATTTGATAGA 61  
 DB 225 LysGluAspProLysSerAlaGlyValAlaThrPheValLleGlnGluPheAspArg 244  
 QY 62 TATTCTGGCTATTGGTGGTGTCCAAAAGCTCAAAACACTCCAGTGGTGTAAATCTTT 121  
 DB 245 PheThrGlyTyrTrpSerProSerAlaValGluAspProAspGlyGlyLysArgVal 264  
 QY 122 AGAATTCATATGAGAAATATGATGATCTGAGGTGGAATATTATTTACATCCCT 181  
 DB 265 TyrLeuLeuTyrGluGluValAspGluThrGluValGluLleIleHisValProSerPro 284  
 QY 182 ATGTGGAACACAGGAGGACAGATTCATTCCTGTTATCTCTAAACAGGTACAGCAATCCT 241  
 DB 285 AlaLeuGluGluArgLysAlaAspAlaTyrArgTyrProArgThrGlySerLysAsnPro 304  
 QY 242 AAAGTCACTTTTAAAGATGTGAGAAATATGATGTGCTGAAGGAGG----- 289  
 DB 305 GlnAlaThrIleLysLeuValGluLleLysThrAspGlnGlnGlyArgValSerLeuCys 324  
 QY 289 ----- 289  
 DB 325 ArgLeuLeuCysValPheLeuLeuArgSerAspArgAlaValLeuThrCysLeuTyr 344



```
RT      initiative";
RL      Dev. Dyn. 225:384-391(2002).
RN      [3]
RP      NUCLEOTIDE SEQUENCE.
RC      TISSUE=Embryo;
RA      Klein S., Gerhard D.S.;
RL      Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC071112; AH71112.1; -, mRNA.
DR      GO; GO:0003824; F: catalytic activity; IEA.
DR      GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR      GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR      InterPro; IPR001375; Peptidase S9.
DR      InterPro; IPR002469; Peptidase S9B.
DR      InterPro; IPR000379; Ser_estr.
DR      Pfam; PF00930; DPPIV_N; 1.
DR      Pfam; PF00326; Peptidase S9; 1.
SQ      SEQUENCE 847 AA; 97484 MW; F515F2609B267BAF CRC64;

Alignment Scores:
Pred. No.:      3,48e-96      Length:      847
Score:          1182.50      Matches:    230
Percent Similarity: 54.0%      Conservative: 58
Best Local Similarity: 43.2%      Mismatches: 68
Query Match:      61.0%      Indels:     177
DB:              2          Gaps:       2

US-10-825-632-8 (1-1083) x Q6GR22_XENLA (1-847)
QY      5 GAAGATGCCAGATCAGCTCGCTACCTTTGTTCTCCAAAGAAGATTGTAGATAT 64
DB      222 ASPASPPCYSSERALA GLYVALALATHR PHEVALILEGLNGLU PHEASPARGPHE 241
QY      65 TCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGCTGGTGTGTAATAATTCTTAGA 124
DB      242 THRGLYTYRTRPSPRPROTHR VALVALASP LEUPROGLYGLYPHE SERVALLEUARG 261
QY      125 ATTCATATGAGAAATGATGAATCTGAGGTGGAAATATTATGTTTACATCCCTATG 184
DB      262 ILELEUTYRGLU GLU VAL ASPGLU SER ASP VAL GLU ILE ILE HIS VAL PROSER PROALA 281
QY      185 TTGGAAACAGGAGGCGACATTTCCTCGTTATCTCTAAACAGGTACAGCAAAATCTCTAAA 244
DB      282 LEU GLU GLU ARG LYS THR ASP ALA TYR ARG TYR PRO ARG THR GLY SER LYS ASN PRO LYS 301
QY      245 GTCACITTTTAAAGATGT CAGAAATAATGATTGATGCTGGAAGAGATCATA----- 295
DB      302 ILETHRLEULYSLEU SERGLU ILE THR THR ASN ASN ASPGLY GLU ILE VAL ALA ALA GLN 321
QY      295 ----- 295
DB      322 ASP LYS GLU LEU VAL LEU PRO PHE HIS SER LEU PHE PRO SER ALA GLU TYR ILE VAL ARG 341
QY      295 ----- 295
DB      342 ALA GLY TRP SER GLN ASP GLY LYS PHE ALA TRP ALA MET LEU LEU ASP ARG PRO GLN CYS 361
QY      295 ----- 295
DB      362 LEU LEU GLN LEU VAL LEU LEU PRO GLU LEU PHE ILE PRO VAL CYS GLU ASP GLU GLU 381
QY      295 ----- 295
DB      382 GLN ARG ILE GLN TYR VAL LYS LYS VAL PRO GLN HIS ILE GLN PRO TYR VAL ILE TYR GLU 401
QY      295 ----- 295
DB      402 GLI LYS SER ASP VAL TRP VAL ASN VAL HIS ASP VAL VAL HIS LEU LEU ARG GLN GLU LYS 421
QY      295 ----- 295
DB      422 GLU ASN GLU ILE ARG PHE ILE CYS SER ILE GLU SER GLN THR GLU PHE CYS HIS LEU TYR 441
QY      295 ----- 295
```

```
Db      442 LEU MET THR SER LEU LEU GLN ASN GLU ARG GLU TRP SER GLN MET THR ILE PRO CYS 461
QY      295 ----- 295
Db      462 THR VAL ASP MET PHE LYS CYS GLN ILE ASN GLN GLU VAL ALA LEU THR SER GLY GLU TRP 481
QY      296 ----- 296
Db      482 GLU VAL LEU CYS ARG HIS GLY ALA LYS VAL TRP VAL ASN GLU ALA MET GLN LEU VAL TYR 501
QY      323 TTTGAAAGGCCAAAGACTCCCTTTAGAGCATACCTGTGACTAGTACGTAGTACGTAAAT 382
Db      502 PHE GLN GLY THR TYR ASP THR PRO LEU GLU HIS LEU TYR VAL THR SER TYR GLN HIS 521
QY      383 CCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCTTGTCTGCTCAGTCAG 442
Db      522 PRO GLY GLU VAL VAL ARG LEU THR GLU ARG GLY TYR SER HIS SER CYS THR MET SER PRO 541
QY      443 CACTGTGACTTCTTTATAAGTATAGTAAACAGAGAATAATCCACACTGTGTGTCCTTT 502
Db      542 ASN PHE ASP MET PHE VAL SER GLN TYR SER VAL SER CYS PRO CYS VAL HIS LEU 561
QY      503 TACAAGCTATCAAGTCTCTGAAAGATGACCCAACTTCCAAAACAAGAGAAATTTTGGCCACC 562
Db      562 TYR ARG LEU ASN GLY ----- ASN PRO LEU TYR GLN HIS PRO GLN PHE TRP ALA SER 578
QY      563 ATTTTGGATTTCACAGGCTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGA 622
Db      579 VAL MET GLU ALA ALA ARG CYS PRO PRO TYR TYR VAL PRO PRO GLU ILE PHE PHE HIS 598
QY      623 AGTACTACTGGATTTACATTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCTGGA 682
Db      599 ALA SER SER GLN VAL LYS LEU TYR GLY MET VAL TYR LYS PRO HIS ASN LEU VAL PRO GLY 618
QY      683 AAGAAATATCTACTGTGCTGTATATATGTTGTTCTCTCAGTCCAGTTCAGTTCGTTGATTAAT 742
Db      619 THR LYS HIS PRO THR VAL LEU PHE VAL TYR GLY TYR GLY PRO GLN VAL GLN LEU VAL ASN ASN 638
QY      743 CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTTTATGTGTT 802
Db      639 SER PHE LYS GLY MET LYS TYR LEU ARG LEU ASN THR LEU ALA HIS LEU GLY TYR ALA VAL 658
QY      803 GTAGTGATACACAAACAGGGGATCTCTGTCACCGAGGCTTAAATTTGAAAGGCGCTTTAAA 862
Db      659 VAL VAL ILE ASP GLY ARG GLY SER CYS HIS ARG GLY LEU ALA PHE GLU GLY ALA LEU LYS 678
QY      863 TATAAATGGGTCAAAATAGAAATTTGACATCAGGTGGAGGACTCCAATATCTAGCTTCT 922
Db      679 ASN LYS MET GLY GLN VAL GLU ILE GLN ASP GLN VAL GLU GLY LEU HIS PHE VAL ALA GLU 698
QY      923 CGATATGATTTTCATTGATTCGTAGTTCGTGGGCATCCAGCTGCTCTATCGAGGATAC 982
Db      699 ARG PHE GLY PHE VAL ASP LEU ASP ARG VAL SER ILE HIS GLY TRP SER TYR GLY PHE 718
QY      982 CTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTCAGGCTGTGCTATTCTCGGGCC 1042
Db      719 LEU SER LEU MET GLY ILE ILE GLN ARG PRO GLU VAL PHE LYS VAL ALA ILE ALA GLY ALA 738
QY      1043 CCAGTCACTCTGTGGATCTTCTTATGATACAGGATACACG 1081
Db      739 PRO VAL THR LEU TRP MET ALA TYR ASP THR GLY TYR THR 751

RESULT 7
O75273 HUMAN
ID O75273 HUMAN PRELIMINARY; PRT; 508 AA.
AC O75273;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE R26984.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN NUCLEOTIDE SEQUENCE.  
 RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,  
 RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,  
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Barnes J.,  
 RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,  
 RA Liu S., Attk C., Andreise T., Frankheim M., Amico-Keller G.,  
 RA Cosfield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,  
 RA Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,  
 RA Kobayashi A., Olsen A.S., Carrano A.V.,  
 RA Submitted (Sep-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC003594; AAC33801.1; -; Genomic DNA.  
 DR MEROPS; S09.019; -;  
 DR Ensembl; ENSG00000142002; Homo sapiens.  
 DR GO; GO:0003824; Fcatalytic activity; IEA.  
 DR InterPro; IPR001375; Peptidase\_S9.  
 DR InterPro; IPR002469; Peptidase\_S9B.  
 DR InterPro; IPR000379; Ser\_estrs.  
 DR Pfam; PF00930; DPPIV\_N; 1.  
 DR Pfam; PF00326; Peptidase\_S9; 1.  
 FT NON TER 1  
 SQ SEQUENCE 508 AA; 57750 MW; 2F4D6645BE2D2C89 CRC64;  
 Alignment Scores:  
 Pred. No.: 5,43e-75 Length: 508  
 Score: 944.00 Matches: 185  
 Percent Similarity: 68.9% Conservative: 48  
 Best Local Similarity: 54.7% Mismatches: 77  
 Query Match: 48.7% Indels: 28  
 DB: 2 Gaps: 5  
 US-10-825-632-8 (1-1083) x 075273\_HUMAN (1-508)  
 QY 125 ATCTATATGAGAAATGATGAATCTGAGTGGAAATATTTCATGTTATCATCCCTATG 184  
 Db 75 valvaltyr-glulgluvalthrAsnValtrpIleAsnValHisAspIlePheTyProPhe 94  
 QY 185 TTGGAACAGAGGAGGCGAGTTCATCCGTTAT-----CCTAAACAGGT 229  
 Db 95 ProGlnSerGluGluGlnSpGluLeuCysPheLeuArgAlaAsnGluCysIlysthrGly 114  
 QY 230 ACNGCAAT---CCTAAAGTCTCT-----TTTAAGATGTCAGAA 265  
 Db 115 PheCysHisLeuTyrrLyseValThrAlaValLeuLysSerGlnGlyTyrrAspTrpSerGlu 134  
 QY 266 ATAATGATGATGCTGAAGGAAG-----ATCATAGTTGATGAAGTC 307  
 Db 135 ProPheSerProGlyGluGlyGluGlnSerLeuThrAsnAlaIleTrpValAsnGluGlu 154  
 QY 308 AGAAGCTGGTATATTTGAAGGACCAAGACCTCCCTTTAGAGCATCACCTGTACGTA 367  
 Db 155 ThrLysLeuValTyrrPheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrrVal 174  
 QY 368 GTCAGTTAGTAAATCCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTC 427  
 Db 175 ValSerTyrrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSer 194  
 QY 428 TGCTGTCATCAGTCAGTCAGTCTCTTTTATAAGTAAAGTATAGTAAACAGGAATCCA 487  
 Db 195 CysSerMetSerGlnAsnPheAspMetPheValSerHisTyrrSerValSerThrPro 214  
 QY 488 CACTGTGTCTCCCTTTACAAGCTATCAAGTCTGAGCATCACCACTTGCACAAACAAG 547  
 Db 215 ProCysValHisValTyrrLysLeuSerGlyProAspAspProLeuHisLysGlnPro 234  
 QY 548 GAATTTGGGCCACCATTTTGGATTCAGAGGCTCTCTCTGACTATCTACTCTCCAGNA 607  
 Db 235 ArgPheTrpAlaSerMetMetGluAlaAla-----Lys 245  
 QY 608 ATTTCTCTTTGAAAGTACTACTGAGTATTCATGTTATGGGATGCTCTACAGGCTCAT 667

Db 246 IlePheHisPheHisThrArgSerAspValArgLeuTyrrGlyMetIleTyrrLysProHis 265  
 QY 668 GATCTACAGCCTGGAAAGAAATATCTCTACGTCTCTCATATATGTTGTTCTCAGGTG 727  
 Db 266 AlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrrGlyGlyProGlnVal 285  
 QY 728 CAGTTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCT 787  
 Db 286 GlnLeuValAsnAsnSerPheLysGlyIleLysTyrrLeuArgLeuAsnThrLeuAlaSer 305  
 QY 788 CTAGGTTATGTGGTTAGTAGATAGACACAGGGGATCTCTGTCACAGGGCTTAAATTT 847  
 Db 306 LeuGlyTyrrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPhe 325  
 QY 848 GAAGGCGCTTTAAATATAAATGGGTCAAATAGAAATTCAGATCAGTGGGAAGGACTC 907  
 Db 326 GluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeu 345  
 QY 908 CAATATCTAGCTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGATCCACGGCTGG 967  
 Db 346 GlnPheValAlaGluLysTyrrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrp 365  
 QY 968 TCCTATGAGGATACCTCTCCCTGAGGCATTAATGCAGAGCTCAGATATCTTCAGGTT 1027  
 Db 366 SerTyrrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysVal 385  
 QY 1028 GCTATTGCTGGGCGCCAGTCACTCTGTGATCTTCTATGATACAGGATACAG 1081  
 Db 386 AlaIleAlaGlyAlaProValThrValTrpMetAlaTyrrAspThrGlyTyrrThr 403  
 RESULT 8  
 QSVCT0 DROME PRELIMINARY; PRT; 1053 AA.  
 ID QSVCT0; Q8IHO7;  
 AC 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
 DE CG3744-PA, isoform A (CG3744-pc, isoform c) (LD3375sp).  
 GN ORFNames=CG3744;  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila;  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Wan K.H., Doyle C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-pfankuch C., Baldwin D.,  
 RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Borchan M.R., Bouck J., Bratkstein P., Brottier P.,  
 RA Burtis K.C., Busan M.R., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Chertys J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delser A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E.I., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fodor C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hartin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Jalali M., Kalush F., Karpen G.H., Kravitz S., Kulp D., Lai Z.,  
 RA Klamel B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.C., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RA "The genome sequence of *Drosophila melanogaster*.";  
RT Science 287:2185-2195(2000).  
RL [2].  
RN  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RA "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
RT *melanogaster* euchromatic genome sequence.";  
RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3].  
RN NUCLEOTIDE SEQUENCE.  
RP MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Swirskas R.,  
RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RT a genome perspective.";  
RT Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
RN [4].  
RN NUCLEOTIDE SEQUENCE.  
RP MEDLINE=22426069; PubMed=12537572;  
RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review.";  
RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [5].  
RN NUCLEOTIDE SEQUENCE.  
RP Berkeley *Drosophila* Genome Project;  
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Friese E., George R.,  
RA Hoskins R., Stapleton M., Pacleb J., Park S., Swirskas R., Smith E.,  
RA Yu C., Rubin G.;  
RT "Drosophila melanogaster release 4 sequence.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6].  
RN NUCLEOTIDE SEQUENCE.  
RP FlyBase;  
RG Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
RN [7].  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=Berkely;  
RA Stapleton M., Brokatein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Drensek D., Farfan D., Friese E.,  
RA George R., Gonzalez M., Guatin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- INTERACTION;  
CC Q9VZ80.CG12016; MbExp=1; IntAct=EBI-105926, EBI-152819;  
CC EMBL; AB003749; AAF56357.2; -; Genomic\_DNA.  
DR EMBL; BT001499; AAN71254.1; -; mRNA.

DR IntAct; Q9VC20; -;  
DR MEROPS; S09.016; -;  
DR Ensembl; CG3744; *Drosophila melanogaster*.  
DR FlyBase; FBGN0039240; CG3744.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001375; Peptidase\_S9.  
DR InterPro; IPR002469; Peptidase\_S9B.  
DR InterPro; IPR000379; Ser\_estr\_1.  
DR Pfam; PF00930; DPPIV\_N; I.  
DR Pfam; PF00326; Peptidase\_S9; 1.  
SQ SEQUENCE 1053 AA; 118053 MW; C94AA663AB464577 CRC64;  
  
Alignment Scores:  
Pred. No.: 7,21e-55 Length: 1053  
Score: 718.50 Matches: 176  
Percent Similarity: 41.1% Conservative: 55  
Best Local Similarity: 31.3% Mismatches: 116  
Query Match: 37.1% Indels: 215  
DB: 2 Gaps: 10  
  
US-10-825-632-8 (1-1083) x Q9VC20\_DROME (1-1053)  
  
QY 5 GAAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCAGAGAAGATTGTGATGATAT 64  
DB 409 AspAspAlaLeuSerAlaGlyValProSerTyrValMetGlnGluGluPheSerArgTyr 428  
QY 65 TCTGCTATTGCTGCTGCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTCCTTACA 124  
DB 429 GlnGlyPheTrpTrpGlnProHis-----SerAsnAspGlyIleTyrArg 443  
QY 125 ATTCTATATGAAGAAATGATGATTCAGTGGAGTGGAAATTTATTCATGTTACATCCCTATG 184  
DB 444 IleValTyrGluGluValAspGluSerGluValSerValTyrThrPheProSerSerThr 463  
QY 185 TTGGAAACAGAGGCGCAGATTCATTCCTTAAACAGGTACACCAATCCTAAA 244  
DB 464 AlaMetHisGlyArgValAspGluTyrArgPheProArgThrGlySerProAsnAlaLys 483  
QY 245 GTCACCTTTTAAAGATG-----TCAGAAATAATG 271  
DB 484 SerLysLeuLysLeuValGlnPheValLeuAsnGluAlaLeuGlnValSerGluIleAla 503  
QY 272 ATT-----GATGCT----- 274  
DB 504 IleLysAspLeuProTyrSerLeuLeuAlaValPheSerTrpLeuGluTyrIleValArg 523  
QY 275 -----GATGCT----- 280  
DB 524 ValGlyTrpThrProAspAlaLysTyrValTrpValGlnGlyLeuAspArgLysGlnGln 543  
QY 280 ----- 280  
DB 544 ArgLeuAspValIleLeuIleProLeuAspAsnPheCysGluSerTyrSerSerGlnVal 563  
QY 280 ----- 280  
DB 564 SerThrProThrAspSerIleGlyAspHisSerTrpArgSerLeuTyrSerArgThrIle 583  
QY 280 ----- 280  
DB 584 ThrProLeuGlnValIleTyrThrGluArgSerAspSerTrpIleAsnValHisAspMet 603  
QY 280 ----- 280  
DB 604 LeuHisPheLeuAspLeuThrGluThrSerValThrPheLeuTrpAlaSerGluGluThr 623  
QY 280 ----- 280  
DB 624 GlyPheArgHisLeuTyrLeuValThrAlaSerLeuLeuLeuSerGlnAlaAsnGlyGln 643





RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Herman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [5];  
 RN NUCLEOTIDE SEQUENCE.  
 RP Berkeley Drosophila Genome Project;  
 RG Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,  
 RA Yu C., Rubin G.;  
 RT "Drosophila melanogaster release 4 sequence.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6];  
 RN NUCLEOTIDE SEQUENCE.  
 RP FlyBase;  
 RG Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AS003749; AAF56356.2; -; Genomic\_DNA.  
 DR MEROPS; S09\_016; -;  
 DR Ensembl; CG3744; Drosophila melanogaster.  
 DR FlyBase; FBgn0039240; CG3744.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001375; Peptidase\_S9.  
 DR InterPro; IPR002469; Peptidase\_S9B.  
 DR InterPro; IPR000379; Ser\_estrs.  
 DR Pfam; PF00930; DPPIV\_N; 1.  
 DR Pfam; PF00326; Peptidase\_S9; 1.  
 SQ SEQUENCE 1113 AA; 123681 MW; 20857E3B212DF2E4 CRC64;

Alignment Scores:  
 Pred. No.: 7.28e-55 Length: 1113  
 Score: 718.50 Matches: 176  
 Percent Similarity: 41.1% Conservative: 55  
 Best Local Similarity: 31.1% Mismatches: 116  
 Query Match: 37.1% Indels: 215  
 DB: 2 Gaps: 10

US-10-825-632-8 (1-1083) x Q9VC19\_DROME (1-1113)

QY	5	GRAGATGCCAGTACGTCGAGTCGCTACCTTTGTTCTCCAGAAAGAAATTTGATAGATAT	64
DB	469	ASpAspAlaLeuSerAlaGlyValProSerTyrValMetGlnGluGluPheSerArgTyr	488
QY	65	TCGTGGTATTGGTGTGTCCAAAGCTCAAACTCCCACTCCCGAGTGGTGTAAATCTTTAGA	124
DB	489	GlnGlyPheTrpTrpGlnProHis-----SerAsnAspGlyIleTyrArg	503
QY	125	ATTCATATGAAGAAATGATGAACTGAGTGGTGAATATTATGTTACATCCCTATG	184
DB	504	IleValTyrGluGluValAspGluSerGluValSerValTyrThrPheProSerSerThr	523
QY	185	TTGAAACAGAGGAGGCAGATTCATTCCGTTATCTCTAAACAGAGTACAGCAATCCCTAAA	244
DB	524	AlaMetHisGlyArgValAspGluTyrArgPheProArgThrGlySerProAsnAlaLys	543
QY	245	GTCACTTTTAAGATG-----TCAGAAATAATG	271
DB	544	SerLysLeuLysLeuValGlnPheValLeuAsnGluAlaLeuGlnValSerGluIleAla	563
QY	272	ATT-----	274
DB	564	IleLysAspLeuProTyrSerLeuLeuAlaValPheSerTrpLeuGluTyrIleValArg	583
QY	275	-----GATGCT-----	280

DB	584	ValGlyTrpThrProAspAlaLysTyrValTrpValGlnGlyLeuAspArgLysGlnGln	603
QY	280	-----	280
DB	604	ArgLeuAspValIleLeuIleProLeuAspAsnPheCysGluSerTyrSerSerGlnVal	623
QY	280	-----	280
DB	624	SerThrProThrAspSerIleGlyAspHisSerTrpArgSerLeuTyrSerArgThrIle	643
QY	280	-----	280
DB	644	ThrProLeuGlnValIleTyrThrGluArgSerAspSerTrpIleAsnValHisAspMet	663
QY	280	-----	280
DB	664	LeuHisPheLeuAspLeuThrGluThrSerValThrPheLeuTrpAlaSerGluGluThr	683
QY	280	-----	280
DB	684	GlyPheArgHisLeuTyrLeuValThrAlaSerLeuLeuSerGlnAlaAsnGlyGln	703
QY	281	-----GAAGGA	286
DB	704	ProAspProGlySerValGlyAlaGlnProSerPheValAspLeuSerAlaLeuGlnPro	723
QY	287	AGGATCATA-----	295
DB	724	ArgIleLeuAsnLysValAlaLeuThrSerGlyGluTrpGluValLeuAlaArgAsnLeu	743
QY	296	---GTTGATGAAGTCAGAAAGCGTGTATATTTTGAAGGCACCAAGACTCCCTTTAGAG	352
DB	744	TrpValAspLysAlaAsnLysLeuValTyrPheValGlyLeuArgAspThrProLeuGlu	763
QY	353	CATCACCCTGACGTACGTACGTACGTAAATCTCTGAGAGGTGACAAAGCTGACGTACCGT	412
DB	764	LysHisLeuTyrValValSerLeuGluArgProGluHisIleArgLeuGluTrpGluPro	783
QY	413	GGCTACTCACATTCCTTGC-----TGCACTCAGT	439
DB	784	GlyTyrSerTyrLeuValGluPheAspAspHisPheAsnAspLeuPheProIlePheSer	803
QY	440	CAGCAGCTGTGACCTTCTTTATAAGTAAGTATAGTAAC---CAGAAAGATCCACACTGTGTG	496
DB	804	GlnGlnCysLysLeuMetLeuLeuValTyrCysAsnIleGlnArgLeuProSerCysLys	823
QY	497	TCCCTTTAAGCTATCAAGTCTCTGAAGATGATGATCCCACTTCCAAACCAAGAAATTTGG	556
DB	824	ValMet-----ArgValAsnGlnThrCysSerAsnGlyGlyValAsn	837
QY	557	GCCACCATTTTGGATTCAGCAGGT-----CCTCTTCTCTGACTAT	595
DB	838	GlyIleGlnIleSerLeuValGlyTyrLeuHisGluGlyGlyLysProGluProGlnTyr	857
QY	596	ACTCCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTTACATTTATGATGGATGTC	655
DB	858	Cys---ProGlnIlePheSerProGlnLeuProSerGlyAspIleValTyrAlaMetVal	876
QY	656	TACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTCTGTCTATATGCT	715
DB	877	PheLysProHisAsnPheGluLeuGlyValLysTyrProThrValLeuAsnValTyrGly	896
QY	716	GGTCCTCAGGTGAGTGGTGAATAATCGTTTAAAGGAGTCAAGTATTTCGCTTGAAT	775
DB	897	GlyProGluValGlnThrValAsnAsnThrPheLysGlyLysHisGlnLeuArgMetHis	916
QY	776	ACCCTAGCTCTCTAGCTTATGTGTTGTAGTGATAGACACACAGGGGATCTCTGTCCCGA	835
DB	917	MetLeuAlaAlaGlnGlyTyrCysValIleCysIleAspSerArgGlySerArgHisArg	936
QY	836	GGGCTTAAATTTGAAGCGCTTTTAAATATAAATCGGTCAATAGAAATTCACCATCAG	895
DB	937	GlyLysArgPheGluSerHisIleArgGlyArgMetGlyGlnValGluLeuThrAspGln	956







```
Db 385 ArgAlaLeuValTyrPheMetGlyLeuArgGluThrProLeuGluGlyHisIleu 404
QY 362 TACGTACTGAGTTTACGTAATCTCGAGAGGTGACAGCGTCACTGCGGTGCTATCA 421
Db 405 TyrValValSerLeuAlaGlnProAsnGlnLeuArgLeuLeuThrMetProGlyTyrSer 424
QY 422 CATCTCTGTCGATCAGTCAGCAGCTGCTCTCTTTAAGTAAATAGTATAGTAAC----- 475
Db 425 -----PheThrValGluPheAsnAspValSer 433
QY 476 -----CAGAAAGTCCACATGCTGTGTCCTTTCAAGCTATCAAGTCTCGAAGTGC 529
Db 434 ValPheGluArgGlyGlyAlaCysProAsnThrAlaAsnGlyCysSerHisGly 453
QY 530 CCACTTGGCAAAACAAGAAATTTTGGCCACCAATTTTGGATTGAGTCAAGAGTCTCTCTCT 599
Db 454 ProThrProProThrProThrProLeuAspAlaLeuArgLeuCysSerValGlyTyrLeuThr 473
QY 590 GACTATATCTCTCT-----CCAGAAATTTTCTCTCTTTTGAAGTACT 628
Db 474 GluGlyGlyProSerGluAsnThrGlnTyrAsnProSerIleHisSerProGlnIleSer 493
QY 629 ACTGGATTTTACATTTGATGGTCTCTACAGGCTCATGATCTACAGCTCGGAAGAA 688
Db 494 SerGlyAspValLeuTyrAlaMetValPheLysProHisAsnPheMetLeuGlyValLys 513
QY 689 TATCTTACTGTGCTGTTCATATATGTTGCTCTCAGGTGCAAGTTTGGTGAATAATCGTTT 748
Db 514 TyrProThrValLeuAsnValTyrGlyGlyProGluValGlnThrValSerAsnThrPhe 533
QY 749 AAGAGGAGTCAAGTATTTCCGCTTGAATACCTACCTGCTCTCTAGGTTATGTTGTTAGTC 808
Db 534 LysGlyMetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyrCysValIleCys 553
QY 809 ATAGACAACAGGGGATCTCTGTCACCGAGGCTTAAATTTGAAGCGCTTTAAATATAA 868
Db 554 ValAspSerArgGlySerArgHisArgGlyValGluPheGluSerTyrIleArgCysArg 573
QY 869 ATGGTCAATGAATTCAGATCAGATGAGTGAAGGATCTCAATATCTAGCTCTTCGATAT 928
Db 574 MetGlyThrValGluLeuSerAspGlnValGluValLeuArgIleLeuAlaAspGlnLeu 593
QY 929 GATTCATTGATTCATGATCGTGTGGGCATCCACCGCTGCTCTATGAGGATACCTCTCTCC 988
Db 594 GlyTyrIleAspMetAspArgValAlaIleHisGlyTyrSerTyrGlyGlyTyrLeuSer 613
QY 989 GTGATGGCATTAAATCAGAGTTCAGATATCTTACGGGTTGCTATTTGCTGGGCCCCAGTC 1048
Db 614 LeuMetGlyLeuValGlnTyrProGluIlePheLysValSerIleAlaGlyAlaProVal 633
QY 1049 ACTCTGTGATCTCTATGATACAGGATACACG 1081
Db 634 ThrSerTrpGluTyrTyrAspThrGlyTyrThr 644

RESULT 12
ID Q7PTT8 ANOGA
AC Q7PTT8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000015447 (Fragment).
GN ORFNames=ENSANGG0000012958;
OS Anopheles gambiae str. PEST;
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCB_TaxID=180454;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC The Anopheles gambiae Sequence Committee;
RG Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
```

```
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008287; EAA03335.3; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_esters-.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 621 621
FT NON_TER 621 621
SQ SEQUENCE 621 AA; 70546 MW; F6EA8463A343BBF3 CRC64;

Alignment Scores:
Pred No.: 4,99e-47 Length: 621
Score: 630.00 Matches: 150
Percent Similarity: 49.6% Conservative: 54
Best Local Similarity: 36.5% Mismatches: 117
Query Match: 32.5% Indels: 90
DB: 2 Gaps: 10

US-10-825-632-8 (1-1083) x Q7PTT8 ANOGA (1-621)
QY 65 TCTGGCTATTGGTGGTGTCCA-----AAAGCTGAACAACACTCCAGTGGTGTAAATTT 118
Db 134 SerSerGlyTyrArgSerProLeuAspLysSerThrThrArgPro----- 148
QY 119 CTGAGAAATCTATATGAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCC 178
Db 149 LeuGlnValIleTyrThrGluThrSerSerSerTyrValAsnValHisAspValLeuGln 168
QY 179 -----CCTATGTTGAAACCAAGGAGGCGAGATTCATTC 211
Db 169 PheValGluLeuSerGluGlnGluValThrPheLeuTyrPalaSerGluGluSerGlyPhe 188
QY 212 CGT-----TATCCTAAAAAGGTACAGCAATCTCT----- 241
Db 189 ArgHisLeuTyrLeuValThrSerSerLeuSerProAsnGluValSerSerGlyAla 208
QY 241 ----- 241
Db 209 GlyAlaThrAspHisSerLeuProSerMetAlaCysIleGlySerThrLeuValAlaArg 228
QY 242 -----AAAGTCATCTTTAAGATGTCAGNAATAATGATGATGCTGAAGNAGG--- 289
Db 229 IleValGlnLysValThrLeuThrGlyAsp-----TrpGluValLeuGlyArgAsn 246
QY 290 ATCATAGTTGATGAGTCAAGTCAAGGCTGTTATTTTGAAGCCACCAAGACTCCCTTTTA 349
Db 247 ValTrpTyrAspArgValArgGlnLeuValTyrPheMetGlyLeuArgGluThrProLeu 266
QY 350 GAGCATCACTGTACGTAGTCACTTAACTCTGAGAGAGGTGACCAAGGCTGACTGAC 409
Db 267 GluLysHisLeuTyrValValSerLeuAlaGlnProAsnGlnLeuArgLeuThrMet 286
QY 410 CGTGGCTACTCATCTTCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTTATAAGTAAGTAT 469
Db 287 ProGlyTyrSerPheThrValGluPheAsnAspCysThrLeuPheLeuGlnThrTyr 306
QY 470 AGTAACACAGAGAATCCACACTGTGTGCTCCCTTTACAAAGCTA-----TCAAGTCCT 520
Db 307 CysAsnIleSerThrLeuProSerTrpGluLeuValArgIleAlaHisAspSerAsnThr 326
```

```
QY 521 GAAGATGACCACTTGCACAAACAAAGGAATTTTGGGCCACCATTGTTGATTTCACGAGT 580
Db      :|||
327 AlaAenGlyAenGlyCys-----SerHisGly 335
QY 581 CCTCTCTGACTATATCTCT-----601
Db      :|||
336 ProThrProPropoThrProIleAspAlaLeuAargLeuCysSerValGlyTyrLeuThr 355
QY 602 -----CCAGAAATTTCTCTTTTGAAGTACT 628
Db      :|||
356 GluGlyGlyProSerGluAenThrGlnTyrAsnProSerIleHisSerProGlnIleSer 375
QY 629 ACTGATTTACATTTGATGGATGCTTACAAGCCTCATGATCTACAGCTCGAAGAAA 688
Db      :|||
376 SerGlyAaspValLeuTyrAlaMetValPheLeuYasProHisAenPheMetLeuGlyVallys 395
QY 689 TATCTCTACTGCTGTTTCATATATGGTGGTCTCAGGTGAGTGTGTAATATCGGTTT 748
Db      :|||
396 TyrProThrValLeuAenValTyrGlyGlyProGluValGlnThrValSerAenThrPhe 415
QY 749 AAAGGAGTCAAGTATTTCCGTTGATATACCTGCTCTAGTGTATGTTGTTGATGTG 808
Db      :|||
416 LysGlyMetAargGlnLeuAargMetHisMetLeuAlaSerGlnGlyTyrCysValIleCys 435
QY 809 ATAGACACAGGGGATCTGTGTCACGGAGGCTTAATTTCAAGCGCCCTTTAAATATAA 868
Db      :|||
436 ValAaspSerArgGlySerArgHisAargGlyValGluPheGluSerTyrIleAargArg 455
QY 869 ATGGGTCAATAGAAATGACGATCAGGTGGAAGGATCCCAATATCTAGCTTCTCGATAT 928
Db      :|||
456 MetGlyThrValGluLeuSerAaspGlnValGluValLeuAargIleLeuAlaAaspGlnLeu 475
QY 929 GATTTTCATGACTAGCTGTGGGCATCCACGGCTGGTCTTATGAGGATACCTCTCC 988
Db      :|||
476 GlyTyrIleAaspMetAaspArgValAlaIleHisGlyTyrSerTyrGlyGlyTyrLeuSer 495
QY 989 CTGATGGCATTAATGAGAGGTGATATCTTACGGGTGCTATTGCTGGGGCCCGGATC 1048
Db      :|||
496 LeuMetGlyLeuValGlnTyrProGluIlePheLysValSerIleAlaGlyAlaProVal 515
QY 1049 ACTCTGTGATCTCTTATGATACAGGATACAG 1081
Db      :|||
516 ThrSerTrpGluTyrTyrAspThrGlyTyrThr 526

RESULT 13
Q54U01.DICDI
ID Q54U01.DICDI PRELIMINARY; PRT; 803 AA.
AC Q54U01; 2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DPB0205566;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.A.,
RA Sugano S., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tungal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero P.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies K., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulseghe H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
```

```
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RT "The genome of the social amoeba Dictyostelium discoideum."
RL Nature 0:0-0(2005).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL: AA01000074; EMBL66699.1; -; Genomic_DNA.
DR Hypothetical protein.
KW SEQUENCE 803 AA; 93187 MW; 338468370C039FAE CRC64;
SQ

Alignment Scores:
Pred. No.: 7,88e-45 Length: 803
Score: 605.50 Matches: 156
Percent Similarity: 39.8% Conservative: 59
Best Local Similarity: 28.9% Mismatches: 122
Query Match: 31.2% Indels: 203
DB: Gaps: 13

US-10-825-632-8 (1-1083) x Q54U01.DICDI (1-803)
QY 14 AGATCAGCTGGAGTCTGCTACCTTTGTTCTCCAGAGAATTTGATAGATATCTGGCTAT 73
Db 189 AtgTyrAlaGlyAaspIleGlyPheIleTyrAlaGluGluPheSerArgTyrThrGlyTyr 208
QY 74 TGTGTGTGTCCTCAAAAGCTGMAACACTCCAGTGTGTGTAATAA---ATTCTTAGAATCTA 130
Db 209 TrpTrpSerProIleValGlyThrCysVallysThrGlyLysProMetTyrThrIleCys 228
QY 131 TATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTGGAA 190
Db 229 TyrLeuGluGluAaspGluThrAsnValMetAaspTyrHisIle-----ProThrSerAsp 246
QY 191 ACAAGG---AGGCGAGATCATTCCTGCTTATCTTAAACAGCGTACAGCAAT--- 238
Db 247 LeuArgGlyLysThrThrGlnTyrLysTyrProLeuAlaGlyLysAsnSerIleCys 266
QY 239 -----CCTAAGTCACCTTTTAAAGATCTCAGAAATA 268
Db 267 LysValCysLeuValSerPheValLeuProThrArgThrThrPheGlnAaspSerLysIle 286
QY 269 ATGATT----- 274
Db 287 GluIleValLysSerGluLeuPheAaspLeuLysThrGlnPheProTrpAlaGluTyrIle 306
QY 274 ----- 274
Db 307 ThrArgAlaGlyTrpThrProAaspGlyHisSerIleTyrLeuGlnLeuLeuAaspArgLys 326
QY 274 ----- 274
Db 327 GlnGlnHisLeuAlaLeuValMetValProLeuHisValPheAlaGluAaspTyrSerSer 346
QY 274 ----- 274
Db 347 SerSerSerSerValLysSerIleProLysLeuProValLeuIleGluGluThrThr 366
QY 275 -----GATGCTGAAGAGGAGTATC 292
Db 367 SerValTrpIleAsnIleGluPheSerPheGlnPheLeuLysSerIleGluAsnGlnLeu 386
QY 293 ATA----- 295
Db 397 IleTrpSerAenGluGlnSerGlyTyrArgHisLeuTyrLeuIleLysTrpAaspLysAen 406
QY 295 ----- 295
Db 407 PheThrAsnIleGlnSerThrThrProIleThrLeuSerThrCysAenAaspAenAasp 426
QY 296 -----GTTGATCAAGTCAAGAGGCTGGTA 319
```

```
Db 427 AsnAenAenTrpMetValSerSerAspAspIleHisIleAspGluLysArgLysVal 446
Qy 320 TATTTTCAAGGACCAAGACCTCCCTTTAGACATCACTCTAGCTAGTACGTTACGTA 379
Db 447 TyrPheThrGlyThrLysAspThrCysLeuGluGlnHisLeuValThrArgPheAsp 466
Qy 380 AATCCT---GGACAGGTGACAGGCTGACTGACCGTGGCTACTCACAT---TCTGTGTC 433
Db 467 LysProAsnSerGluIleLysArgLeuSerHisAlaAsnPheSerHisArgSerIleSer 486
Qy 434 ATCAGTCAGACGTCGACTCTCTTTATAGTAAGTATAGTAAACCAAGAAAT--- 484
Db 487 IleSerSerAsnPhelyLysPheIleThrThrTyrSerAsnIleSerThrIleSerLys 506
Qy 485 -----CCA 487
Db 507 ThrGluValPheAspLeuIleTyrAsnAspAspAsnAsnAspAsnAspIleTyrPro 526
Qy 488 CACTGTGTCTCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCACAAACAAAG 547
Db 527 IleValLysSerSerPhePheIleAsnAspAspAspAspAspAspAspAspLysLys 546
Qy 548 GAATTTTGGGCCACCATTTTGGATTTCAGCAGGCTCTCTCTGACTATCTCTCCAGAA 607
Db 547 LysIleAsnIleAsnIle-----ProLys 554
Qy 608 ATTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTACAGCCCTCAT 667
Db 555 IlePheAsnPhelyLysAsnSerLysGlyValThrIleTyrGlyGlnTyrThrLeuProSer 574
Qy 668 GATCTACAGCCTGGAAGAAATATCTACTGCTGCTTCATATATATATATATATATATATAT 727
Db 575 AspTyrSerLysAspLysLysTyrProThrValValTyrValTyrGlyGlyProHisVal 594
Qy 728 CAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGATACCTAGCCTCT 787
Db 595 GlnIleValArgAsnGlnTyrAsnTyrIleLysGlnHisTyrThrAsn----- 610
Qy 788 CTAGGTTATGTTGTAGTATAGACACAGAGGGATCTCTGTCACCGAGGCTTAAATTT 847
Db 611 PheGlyPheIleGlnValMetIleAspAsnValGlySerAlaAsnArgGlyLeuGluPhe 630
Qy 848 GAAGGCGCTTTAAATATATAATGCTCAATAGAAATTCACATCAGGTGGAAGGACTC 907
Db 631 GluSerHisIleArgGluLysMetGlyGlnValGluIleGlyAspGlnValGluGlyIle 650
Qy 908 CAATATCAGCTCTCGATATGATTC-----ATTGACTTATGATCGTGGCCATCCAC 961
Db 651 AsnTyrLeuValGlyAsn---AspIleValSerIleAspValAsnArgIleAlaIleSer 669
Qy 962 GCCTGCTCTATGAGATACCTCTCCCTGATGCAATTAATGACAGGTCAGATATCTTC 1021
Db 670 GlyTrpSerTyrGlyTyrAsnSerLeuMetAlaIleSerGlnArgProAspValPhe 689
Qy 1022 AGGTTGCTATTGCTGGGCCCCAGTCACCTCTGTGATCTCTATGATACAGGATACACG 1081
Db 690 LysIleAlaValCysGlyAlaProValSerAspTrpArgLeuTyrAsnThrGlyTyrThr 709

RESULT 14
Q5TXJ2 ANOGA PRELIMINARY; PRT; 557 AA.
AC Q5TXJ2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE ENSANGP0000029249 (Fragment)
GN ORFNames=ENSANGG0000012958;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
```

```
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008287; EAL42106.1; -; Genomic_DNA.
DR GO; GO:0003824; Fcatalytic activity; IEA.
DR GO; GO:0008236; Fserine-type peptidase activity; IEA.
DR GO; GO:0006508; Pproteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9.
DR InterPro; IPR000379; Ser_pstrs.
DR Pfam; PF00930; DPEIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 557
SQ SEQUENCE 557 AA; 62685 MW; 08AAA0944ACD0808 CRC64;
```

## Alignment Scores:

Pred. No.:	1.69e-44	Length:	557
Score:	601.50	Matches:	129
Percent Similarity:	63.6%	Conservative:	42
Best Local Similarity:	48.0%	Mismatches:	79
Query Match:	31.0%	Indels:	19
DB:	2	Gaps:	5

US-10-825-632-8 (1-1083) x Q5TXJ2\_ANOGA (1-557)

```
Qy 299 GATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGACATCAC 358
Db 231 AspArgValArgGlnLeuValTyrPheMetGlyLeuArgGluThrProLeuGluLysHis 250
Qy 359 CTGTAGTAGTCAGTACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTAC 418
Db 251 LeuTyrValValSerLeuAlaGlnProAsnGlnLeuArgLeuLeuThrMetProGlyTyr 270
Qy 419 TCACATTTCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTT---ATAAGTAAGTATAGTAC 475
Db 271 SerPheThrVal-----GluPheAsnAspLeuValArgIleAlaHisAspSerAsn 287
Qy 476 CAGAGAATCCACACTGTGTGTCCTTTTACAGCTATCAAGCTATCAAGATGACCACT 535
Db 288 ThrAlaAsnGlyAsnGlyCysSer-----HisGlyProThrProProProThr 303
Qy 536 TGCACAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGCTCTCTTCTGACTAT 595
Db 304 -----ProIleAspAlaLeuArgLeuCysSerValGlyTyrLeuThrGluGly 319
Qy 596 ACTCCT-----CCAGAAATTTCTCTTTTGAAGTACTACTGGA 634
Db 320 GlyProSerGluAsnThrGlnTyrAsnProSerIleHisSerProGlnIleSerSerGly 339
Qy 635 TTTACATTGATGGATGCTCTACAGCCTCATGCTACAGCCTGGAAGAAATATCTCT 694
Db 340 AspValLeuTyrAlaMetValPheLysProHisAsnPheMetLeuGlyValLysTyrPro 359
Qy 695 ACTGCTGTTTCATATATATGTTGTTCTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 754
Db 360 ThrValLeuAsnValTyrGlyProGluValGlnThrValSerAsnThrPheLysGly 379
Qy 755 GTCAGATTTTCCGCTTGAATACCTGACTCTCTAGGTTATGTTGTTGTTGTTGTTGTTGTTGTT 814
Db 380 MetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyrCysValIleCysValAsp 399
```

```
Qy 815 AACAGGGATCTGTCCAGGAGGCTTAAATTTGAAGGCGCTTTAAATATAAATGGGT 874
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
400 SerArgglySerArgHisArgGlyValGluPheGluSerTyrIleArgArgMetGly 419
Qy 875 CAAATGAATTCACGATCAGGTGGAGGAGCTCCCAATATCTAGCTTCTCGGATATGATTC 934
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
420 ThrValGluLeuSerAspGlnValGluValLeuArgIleLeuAlaAspGlnLeuGlyTyr 439
Qy 935 ATTGAATTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATG 994
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
440 IleAspMetAspArgValAlaIleHisGlyTyrSerTyrGlyGlyTyrLeuSerLeuMet 459
Qy 995 GCATTATGAGAGGTCAGATATCTTCAGGTTTGTCTATGCTGGGGCCCAAGTCACCTG 1054
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
460 GlyLeuValGlnTyrProGluIlePheLysValSerIleAlaGlyAlaProValThrSer 479
Qy 1055 TGGATCTTCATGATACAGGATACAG 1081
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
480 TrpGluTyrThrAspThrGlyTyrThr 488

RESULT 15
O9A6E0 CAUCR
ID O9A6E0 CAUCR PRELIMINARY; PRT; 738 AA.
AC O9A6E0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DI Dipeptidyl peptidase IV.
GN OrderedLocusNames=CC2154;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Uitterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005887; AAK24125.1; -; Genomic_DNA.
DR PIR; A87516; A87516.
DR TIGR; CC2154; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser. esters.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
DR KW complete proteome.
SQ SEQUENCE 738 AA; 80591 MW; 1AA8E8192164F0EA CRC64;

Alignment Scores:
Pred. No.: 6 82e-38 Length: 738
Score: 527.50 Matches: 137
Percent Similarity: 41.2% Conservative: 61
Best Local Similarity: 28.5% Mismatches: 132
Query Match: 27.2% Indels: 151
DB: Gaps: 13

US-10-825-632-8 (1-1083) x O9A6E0_CAUCR (1-738)
Qy 5 GAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAGAGAAGATTTGATAGATAT 64
```

```
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
189 LysAspAlaLeuSerPheGlyValAlaGluPheIleValGlnGluLeuAspArgPhe 208
Qy 65 TCTGCTATTGCTGCTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGA 124
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
209 ThrGlyTyrTrpTyrSerProAspGluSer-----Arg 219
Qy 125 ATTCTATATGAAGAAATATGATCTGAGGTGGAATATTATCATGTTTACATCCCTATG 184
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
220 IleValTyrThrArgValAspGluSerGlyValAspIleVal-----233
Qy 185 TTGGAACACAGGAGGCGAGAT-----TCATTCCGTTAT 217
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
234 -----ProArgAlaAspIleGlyProGlyGlyAlaThrValValAsnGlnArgTyr 250
Qy 218 CTTAAACAGCATACAGCAAAATCCT-----241
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
251 ProArgAlaGlyArgProAsnAlaValAspLeuPheValArgAspLeuAlaSerGly 270
Qy 242 AAGTCACATTT-----253
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
271 LysValThrAlaLeuAspLeuGlyAlaAsnLysAspIleTyrValAlaArgValAlaTrp 290
Qy 254 -----AAGATGTCAGAA-----265
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
291 SerAlaAspGlyLysThrValTyrValGlnArgLeuSerArgAspGlnLysThrLeuAsp 310
Qy 266 -----ATAATGATTCAT-----277
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
311 LeuLeuAlaPheAspAlaAlaThrGlyAlaGlyLysThrIleLeuThrAspThrAspPro 330
Qy 277 -----277
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
331 HisPheIleGluValSerAsnAspPheArgProLeuThrAspGlyThrPheLeuTrpGly 350
Qy 278 -----GCTCAAGGAAGGATCATATA-----295
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
351 SerGluLysAspGlyAsnGlnHisLeuTyrArgTyrAlaAlaAspGlyLysLeuIleAla 370
Qy 296 -----GTTGATCAAGTCAGAGG 313
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
371 GlnIleThrLysGlyAspTyrProValIleGlyLeuGluGlyValAspGluAlaArgLys 390
Qy 314 CTGGTATATTTTGAAGGCACAAAGACTCCCTTTAGAGCATCACCCTGTACGTAGTCAGT 373
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
391 ValAlaIlePheSerAlaSerIleAspThrProIleGluArgLeuTyrGluValSer 410
Qy 374 TACGTAAATCCTGGAGAGTGACCAAGCGTGACCGGTGGTACTCATCATTTCTTGGCTGC 433
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
411 TyrAlaLysProGlyLysProLysAlaLeuThrSerAlaGlyGlyTyrTrpAlaAlaLys 430
Qy 434 ATCAGTCAGCACTGTGACTTCTTTTATAGTAACTAGTAAACAGAGATCCACACTGT 493
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
431 ValAlaAspAsnGlyGlyAlaPheAlaGlyThrTyrSerAspProLysThrProSerGln 450
Qy 494 GTGTCCTTTTACAAAGCTATCAAGTCCT-----GAAGATCACCAACTTGC 538
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
451 ThrAlaLeuTyrSerAlaAspGlyLysArgValArgTyrIleGluGluAsnLysLeuAla 470
Qy 539 AAAACAAAGGAATTTTGGCCACCATTTTGGATTTCAGAGCGCTCTTCTGACTATACT 598
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
471 GluGlyHisProTyrTrp-----ProTyrAlaAlaAsnLeuProGln-----484
Qy 599 CTTCCAGAAATTTTCTCTTTTGAAGTACTTACCTGGATTTTACATTTGATGGGTGCTTAC 658
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
485 ---ProGluPheGlySerLeuLysAlaAlaAspGlyGluThrLeuHisTyrGluIleLeu 503
Qy 659 AAGCTCATGATCTACAGCCTGGAAAGAAATATCTTACTGTCTGTTCATATATGCTG 718
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
504 LysProIleGlyPheAspProAlaLysTyrProAlaIleValSerValTyrGlyGly 523
Qy 719 CTTAGGTGCGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACC 778
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
```

```

Db      524 ProHisAlaGlnargValMetLysAenTrpHisSerProSerGlu-----ArgThr 540
Qy      779 CTAGCCTCTCTAGTTATGTGTTGTAGTAGACACAGGGGATCTCTGCACCGAGG 838
Db      541 TyrLeuGluAlaGlyTyrValIlePheLysLeuAspAsnargGlySerGlyAsnargSer 560
Qy      839 CTTAAATTTGAAGCGCCTTTAAATATATAAATGGTCAAATAGAAATTGACGATCAGGTG 898
Db      561 AlalysPheMetArgAlaLeuAspArgLysLeuGlyThrValGluValGluAspGlnLeu 580
Qy      899 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTGACTTAGATCGTGTGGGCATC 958
Db      581 LeuGlyAlalysPheLeuAlaSerGln---ProTyrValAspAlaAspLysLeuGlyVal 599
Qy      959 CACGGCTGTCCTATGGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGTCAGATATC 1018
Db      600 MetGlyTrpSerTyrGlyPheMetAlaLeuMetLeuLeuThrAlaGluAsnThrPro 619
Qy      1019 TTCAGGGTTGCTATTGCTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATAC 1078
Db      620 PheLysAlaGlyAlaAlaGlyAlaProProThrGluTrpSerLeuTyrAspThrAlaTyr 639
Qy      1079 ACG 1081
Db      640 Thr 640

```

Search completed: May 2, 2006, 01:34:05  
 Job time : 274.427 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 01:34:19 ; Search time 6.63677 Seconds  
(without alignments)  
2698.234 Million cell updates/sec

Title: US-10-825-632-8

Perfect score: 1938

Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatacacgga 1083

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/abs/ABSWEB.spool/US10825632/runat 01052006 105950 3311/app query.fasta\_1  
-DB=Issued\_Patents\_AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs802p -USER=US10825632@cgn\_1\_131@runat 01052006 105950 3311  
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5 COMB.pcp:\*  
2: /cgn2\_6/ptodata/1/iaa/6 COMB.pcp:\*  
3: /cgn2\_6/ptodata/1/iaa/H COMB.pcp:\*  
4: /cgn2\_6/ptodata/1/iaa/PCUS COMB.pcp:\*  
5: /cgn2\_6/ptodata/1/iaa/RE COMB.pcp:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	1933	99.7	360	2	US-10-070-464-7	Sequence 7, Appli
2	1836.5	94.8	882	2	US-09-976-674-1	Sequence 1, Appli
3	1836.5	94.8	882	2	US-10-070-464-1	Sequence 7, Appli
4	1226	63.3	690	2	US-09-976-674-7	Sequence 3, Appli
5	1220.5	63.0	863	2	US-09-976-674-3	Sequence 23, Appl
6	1220.5	63.0	892	2	US-09-976-674-23	Sequence 27, Appl
7	1220.5	63.0	892	2	US-09-976-674-27	Sequence 11, Appl
8	1209	62.4	661	2	US-09-976-674-11	Sequence 33, Appl
9	1201	62.0	879	2	US-09-976-674-33	Sequence 19, Appl
10	1201	62.0	879	2	US-09-976-674-35	Sequence 19, Appl
11	1199.5	61.9	658	2	US-09-976-674-19	Sequence 23, Appl
12	1132.5	58.4	832	2	US-09-976-674-29	Sequence 31, Appl

13	1132.5	58.4	832	2	US-09-976-674-31	Sequence 37, Appl
14	1113	57.4	819	2	US-09-976-674-37	Sequence 39, Appl
15	1113	57.4	819	2	US-09-976-674-39	Sequence 4, Appli
16	1086.5	56.1	310	2	US-09-794-236-4	Sequence 3, Appli
17	1086.5	56.1	310	2	US-10-070-464-3	Sequence 21, Appl
18	931.5	48.1	613	2	US-09-976-674-21	Sequence 5, Appli
19	860.5	44.4	465	2	US-10-070-464-5	Sequence 13, Appl
20	584	30.1	358	2	US-09-976-674-13	Sequence 16812, A
21	444	22.9	981	2	US-09-902-540-16812	Sequence 32313, A
22	403.5	20.8	159	2	US-09-270-767-32313	Sequence 30, Appl
23	403.5	20.8	732	2	US-09-518-550-30	Sequence 28, Appl
24	369	19.0	710	2	US-09-518-550-28	Sequence 25, Appl
25	352	18.2	518	2	US-09-976-674-25	Sequence 2, Appli
26	347.5	17.9	771	2	US-08-462-284-2	Sequence 2, Appli
27	347.5	17.9	771	2	US-09-079-592-2	Sequence 42, Appli
28	344	17.8	676	2	US-09-518-550-42	Sequence 29, Appl
29	344	17.8	723	2	US-09-518-550-29	Sequence 11421, A
30	329	17.0	737	2	US-09-902-540-11421	Sequence 3, Appli
31	315	16.3	755	4	PCT-US93-07923-3	Sequence 2, Appli
32	315	16.3	759	4	PCT-US93-07923-2	Sequence 3, Appli
33	315	16.3	766	1	US-08-230-491A-3	Sequence 3, Appli
34	315	16.3	766	1	US-08-619-280A-3	Sequence 3, Appli
35	315	16.3	766	1	US-08-940-391-3	Sequence 1, Appli
36	315	16.3	766	2	US-09-794-236-1	Sequence 6, Appli
37	315	16.3	766	2	US-10-002-593-6	Sequence 6146, Ap
38	315	16.3	766	2	US-09-949-016-6146	Sequence 3, Appli
39	315	16.3	766	2	US-09-265-606-3	Sequence 6, Appli
40	315	16.3	766	2	US-10-423-714-6	Sequence 10450, A
41	315	16.3	775	2	US-09-949-016-10450	Sequence 27, Appl
42	309	15.9	766	2	US-09-518-550-27	Sequence 11805, A
43	287.5	14.8	574	2	US-09-902-540-11805	Sequence 3, Appli
44	280.5	14.5	818	2	US-09-462-845-3	Sequence 3, Appli
45	280.5	14.5	818	2	US-10-402-312-3	

#### ALIGNMENTS

##### RESULT 1

US-10-070-464-7  
; Sequence 7, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: GH-007  
; CURRENT APPLICATION NUMBER: US/10/070,464  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-070-464-7

Alignment Scores:  
Pred. No.: 3.13e-220 Length: 360  
Score: 1933.00 Matches: 360  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 99.7% Indels: 0  
DB: 2 Gaps: 0

US-10-825-632-8 (1-1083) x US-10-070-464-7 (1-360)

QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTTGTTCCTCCAAAGAAGATTGTGATAGA 61

```
Db 1 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluPheAspArg 20
Qy 62 TATTCTGGCTATTGGTGTCTCCAAAGCTGAACCACTCCAGTGGTGTAAATTCCT 121
Db 21 TyrSerGlyTyrTrpCysProLysAlaGluThrProSerGlyGlyLysLeu 40
Qy 122 AGAATCTATATGAAGAAATGATCAATCTGAGGTGGAAATATTATCATGTTACATCCCT 181
Db 41 ArgLeuLeuTyrGluGluAsnAspGluSerGluValGluLeuIleHisValThrSerPro 60
Qy 182 ATCTGGACACAGGAGGACATTCATCTCCGTTATCTCTAAACAGGTACAGCAATTCCT 241
Db 61 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 80
Qy 242 AAGTCACTTTTAAGTATCAGAAATAATGATGTGCTGAAGGAAGGATCATAGTTGAT 301
Db 81 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleValAsp 100
Qy 302 GAAGTCAGAGGCTGTATATTTTGAAGGCCCAAGACTCCCTTTAGAGCATCACCTG 361
Db 101 GluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeu 120
Qy 362 TAGCTAGTCAGTTACGTAATCCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCA 421
Db 121 TyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSer 140
Qy 422 CATCTCTGCTGATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGTACCAAG 481
Db 141 HisSerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGluLys 160
Qy 482 AATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGTACCACTGACCA 541
Db 161 AsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 180
Qy 542 ACAAGGAATTTTGGGCCACCATTTTGATTCACAGCAGTCTCTCTCTGACTATCTCT 601
Db 181 ThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrPro 200
Qy 602 CCAGAAATTTCTCTTTTGAAGTACTGGAATATTCATTGATGGGATGCTCTACAG 661
Db 201 ProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLys 220
Qy 662 CCTCATGCTCTACAGCTGGAGAAATATCCTTACTGTGCTGTCTTATATATGCTGCT 721
Db 221 ProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyPro 240
Qy 722 CAGGTGCAAGTTCGTAATATCGGTTTAAAGGAGTCAAGTATTCGCTTGAATACCTTA 781
Db 241 GlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeu 260
Qy 782 GCCTCTAGGTTAGTGTGTAGTATAGACAAACAGGGGATCTCTGTCCAGCGGGCTT 841
Db 261 AlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeu 280
Qy 842 AAATTGAGCGGCTTTAAATATAAATGGGTCAAAATAGAAATGACCATCAGGTGAA 901
Db 281 LysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGlu 300
Qy 902 GAGCTCAATATCTAGCTTCTCGATATGATTTCATTGACTATGATCGTGGTGGCATCCAC 961
Db 301 GlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyLeHis 320
Qy 962 GCCTGGTCTATGAGGATACCTCTCCGTGATGCAATTAATCAGAGGTTCAGATATCTTC 1021
Db 321 GlyTrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePhe 340
Qy 1022 AGGGTTGCTATTGCTGGGGCCAGTCATCTCTGTGATCTTCTATGATACAGGATACAG 1081
Db 341 ArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 360
```

RESULT 2

US-09-976-674-1

```
; Sequence 1, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-1
Alignment Scores:
Pred. No.: 1,369-208 Length: 882
Score: 1826.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 2 Gaps: 1
US-10-825-632-8 (1-1083) x US-09-976-674-1 (1-882)
Qy 2 GAAGAAGATCCAGATCAGCTGAGTCCGCTACCTTTGTTCTCCAAAGAAGATTGTAGAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluPheAspArg 263
Qy 62 TATTCTGGCTATTGGTGTGTCTCCAAAGCTGAACCACTCCAGTGGTGTAAATTCCT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrProSerGlyGlyLysLeu 283
Qy 122 AGAATCTATATGAAGAAATGATCAATCTGAGGTGGAAATATTATCATGTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
Qy 182 ATGTTGGAAACAAAGGAGGCGAGATTCATTCGTTATCTCTAAACAGGTACAGCAATTCCT 241
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
Qy 242 AAAGTCACCTTTTAAAGATGTCAGAAATAATGATTCATCTGAGGTGGAAATATTATCATGTTACATCCCT 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
Qy 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
Qy 295 ----- 295
Db 364 ArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
Qy 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
Qy 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
Qy 295 ----- 295
Db 424 GluGluThrThrAspIleIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
Qy 295 ----- 295
```

444 HisGluGluGluLeuPheAlaSerGluCysLysThrGlyPheArgHisLeu 463  
295 -----  
464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483  
295 -----  
484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503  
296 -----GTTGATGAAGTCAGAGCGCTGATAT 322  
504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523  
323 TTTGAGGACCAAGACCTCCCTTTAGAGCATACCTGTAGCTAGTCAGTTACGTAAAT 382  
524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543  
383 CCTGGAGAGTGACAGGCTGACTACCGTGCTACTCACATTCTTGTGTCATCAGTCAG 442  
544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563  
443 CACTGTGACTCTTTATAGTAGTATAGTACCAAGAGAAATCCACACTGTGTGCTCCTT 502  
564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
503 TACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGAAATTTGGCCAC 562  
584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
563 ATTTTGAATCAGAGCTCTCTCTGACTATACCTCCAGAAATTTCTCTTTTGAA 622  
604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
623 AGTACTAGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTCGA 682  
624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProgly 643  
683 AAGAAATATCTACTGCTGTTTCATATATGTTGGTCTCCTCAGTGCAGTTGGTGAATAT 742  
644 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAsn 663  
743 CGTTTAAAGCAGTCAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTTATGTGTT 802  
664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
803 GTAGTAGACACACAGGGGCTCCTGTACCGAGGGCTTAAATTTGAAGCGCCCTTTAAA 862  
684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
863 TATAAATGGGTCAAAATGAATTTGACGATCAGTGGAGGACTCCAATATCTAGCTTCT 922  
704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
923 CGATATGATTTCAATGACTAGATCGTGTGGGCAATCCACGCTGCTGCTTATGGAGATAC 982  
724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743  
983 CTCCTCCGATGGCAATTAATGACAGGTCAGATATCTTCAGGCTGCTATTTGCTGGGGCC 1042  
744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
1043 CCAGTCACCTGTGGATCTTCTATCATACAGATACACG 1081  
764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

RESULT 3  
US-10-070-464-1  
; Sequence 1, Application US/10070464  
; Patent No. 6881564  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: GORRELL, Mark Douglas

TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
FILE REFERENCE: GH-007  
CURRENT APPLICATION NUMBER: US/10/070,464  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: PCT/AU00/01085  
PRIOR FILING DATE: 2000-09-11  
PRIOR APPLICATION NUMBER: AU PQ5709  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: AU PQ2762  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 882  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-070-464-1

Alignment Scores:  
Pred. No.: 1.36e-208 Length: 882  
Score: 1836.50 Matches: 360  
Percent Similarity: 67.5% Conservative: 0  
Best Local Similarity: 67.5% Mismatches: 0  
Query Match: 94.8% Indels: 173  
DB: 2 Gaps: 1

US-10-825-632-8 (1-1083) x US-10-070-464-1 (1-882)

QY 2 GAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAGAGAATTTGTAGTA 61  
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263  
QY 62 TATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATCTT 121  
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283  
QY 122 AGAATCTATATGAAGAAATGATGATCTGAGTGGAAATTTATCATGTTACATCCCT 181  
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303  
QY 182 ATGTTGGAAACAGGAGGCGAGATTCATTCGTTTATCTTAAACAGGTACAGCAATCCT 241  
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323  
QY 242 AAAGTCATTTTAAAGATGTCAGAAATATGATGATGCTGCTGAAGGAGGATCATA---- 295  
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343  
QY 295 ----- 295  
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363  
QY 295 ----- 295  
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383  
QY 295 ----- 295  
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403  
QY 295 ----- 295  
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423  
QY 295 ----- 295  
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443  
QY 295 ----- 295  
Db 444 HisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463  
QY 295 ----- 295

```
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuPro 483
Qy 295 -----
Db 484 AlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSerGlyGluTrp 503
Qy 296 -----
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523
Qy 323 TTGAAGGACCAAGAGATCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSerTyrValAsn 543
Qy 383 CTGGAGAGGTCAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGCATCGTAC 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
Qy 443 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCCCTT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 503 TACAAGCTATCAAGTCCCTGAAGATGACCCAACTTGCACAAACAAAGAAATTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
Qy 563 ATTTGGGATTCAGCAGGTCCCTTCTCCTGACTATACCTCCAGAAATTTCTTTTGA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
Qy 623 AGTACTACTGATTTACATGATGGATGCTCTACAGCTCATGATCTACAGCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 683 AAGAAATATCTACTGCTGCTTCATATATGTTGGTCTCTCAGGTACGTTGGTGAATAAT 742
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 743 CGGTTTAAAGAGTCAAGTATTTCCCTTGAATACCTTAGCTCTCTAGGTTATGTGTT 802
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 803 GTAGTGTATAGACACAGGGGATCTGTCCAGGAGGCTTAAATTTGAAGGCGCTTTAAA 862
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
Qy 863 TATAAATGGTCAAAATAGAAATTCAGCATCAGGTGGAGGACTCCCAATATCTAGCTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
Qy 923 CGATATGATTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGAGGATAC 982
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyrGlyTyr 743
Qy 983 CTCTCCCTGATGGCATTATGACAGAGGTACAGATATCTTCAGGTTGCTATGCTGGGCC 1042
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
Qy 1043 CCAGTCACTCTGTGATCTTCTATGATACAGGATACACG 1081
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776
```

## RESULT 4

```
US-09-976-674-7
; Sequence 7, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
```

```
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-7
Alignment Scores:
Pred. No.: 3,46e-136 Length: 690
Score: 1226.00 Matches: 259
Percent Similarity: 48.5% Conservative: 0
Best Local Similarity: 48.5% Mismatches: 0
Query Match: 63.3% Indels: 275
DB: 2 Gaps: 2
US-10-825-632-8 (1-1083) x US-09-976-674-7 (1-690)
Qy 2 GAAGAAGATGCCAGATCAGCTGAGTGGCTACTTCTTCTCCAAAGAAATTTGTAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
Qy 62 TATTCCTGCTATTGGTGTCTCCAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTCCT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
Qy 122 AGAATTCCTATGAAGAAATGATGAATCTAGGTGGAAATTTATTCATGTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluValGluIleIleHisValThrSerPro 303
Qy 182 ATGTTGGAAACAGAGGGGAGATTCATTCCTGTTATCTTAAACAGGTACAGCAATCT 241
Db 304 MetLeuGlnThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
Qy 242 AAATCTCTTTAAGATGTCAGAAATTAATGATGATGCTGAAGGAGGATCATA----- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
Qy 295 -----
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
Qy 295 -----
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
Qy 295 -----
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
Qy 295 -----
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
Qy 295 -----
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
Qy 295 -----
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
Qy 295 -----
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuPro 483
Qy 295 -----
```

484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503  
296 -----GTTGATGAAGTCAGAGCGCTGTATAT 322  
504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523  
323 TTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAAT 382  
524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuValValValSerTyrValAsn 543  
383 CCTGGAGGTGACCAAGGTGACTGACCGTGGCTACTCACATTCTTGTGTCATCAGTCAG 442  
544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563  
443 CACTGTGACTCTTTTAAAGTAAAGTAAACCAAGAGATCCACACTGTGTCTCCTT 502  
564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
503 TACAAGCTATCAAGTCTGAGATGACCAACTTCGCAAAACAAAGGAATTTTGGCCACC 562  
584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheThrPalaThr 603  
563 ATTGTGATTCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTTCTCTTTGAA 622  
604 IleuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
623 AGTACTACTGATTTACATTGTATGGATGCTCTACAGCCTCATGATCAGCCTGGA 682  
624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
683 AAGAAATCTCTACTGCTGCTCATATATGTTGGTCTCCTCAGTGCAGTGTGTGAATAAT 742  
644 LysLysTyrProThrValLeuPheIleTyrGlyGly ----- 655  
743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTTATGTGTT 802  
655 ----- 655  
803 CTAGTGATAGACAAAGGGATCTGTCTACGGAGGGCTTAATTTGAAGCGCCTTTAA 862  
655 ----- 655  
863 TATAAATGGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922  
655 ----- 655  
923 CGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGCTTATGAGGATAC 982  
655 ----- 655  
983 CTCTCCCTGATGCCAATTAATGACAGGTACAGATATCTCAGGCTTGTCTTGTGGGCC 1042  
656 -----Arg-LeuLeuLeuLeuGlyPr 662  
1043 CAGTCACTCTGATCTCTATGATACAGGATACACGG 1082  
662 oGlnSerLeuCysGlySerSerMetIleGlnAspThrArg 675

RESULT 5  
US-09-976-674-3  
; Sequence 3, Application US/09976674  
; Patent No. 684180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 863  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-3  
Alignment Scores:  
Pred. No.: 1,73e-135 Length: 863  
Score: 1220.50 Matches: 236  
Percent Similarity: 54.3% Conservative: 54  
Best Local Similarity: 44.2% Mismatches: 69  
Query Match: 63.0% Indels: 175  
DB: 2 Gaps: 2  
US-10-825-632-8 (1-1083) x US-09-976-674-3 (1-863)  
QY 5 GAAGATGCCAGTCAGCTGAGTCGCTTGTCTCCAAAGAAATTTTCATAGATAT 64  
Db 234 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 253  
QY 65 TCTGCTATTGCTGTGTCCAAAGCTGAAACCAACTCCCAGTGGTGGT --- AAAATCTTT 121  
Db 254 ThrGlyTyrTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 273  
QY 122 AGAATTTCTATATGAAGAAATATGATCTGAGTGGAAATTTATTTATTTATCTTACATCCCT 181  
Db 274 ArgIleLeuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro 293  
QY 182 ATGTTGGAAACACAGGAGGCGCAGATTCATTCGTTTATCTTAAACAGGATCAGCAATCCT 241  
Db 294 AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 313  
QY 242 AAAGTCACTTTTAAAGATCTCAGAAATAATGATTCATCTGAAAGAGGATCATA --- 295  
Db 314 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 333  
QY 295 ----- 295  
Db 334 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 353  
QY 295 ----- 295  
Db 354 ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 373  
QY 295 ----- 295  
Db 374 GlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIleProSerThrGluAsnGlu 393  
QY 295 ----- 295  
Db 394 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr 413  
QY 295 ----- 295  
Db 414 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 433  
QY 295 ----- 295  
Db 434 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 453  
QY 295 ----- 295  
Db 454 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 473  
QY 295 ----- 295  
Db 474 ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu 493  
QY 296 -----GTTGATGAAGTCAGAGCGCTGTATA 319  
Db 494 TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuVal 513

```
OY 320 TATTTCAGGACCAAGGCTCCCTTTAGAGCATCAGCTGTACGTAGTCAGTTAGTA 379
Db 514 TyPheGlnGlyThrLysAspThrProLeuGluHisLeuTyValValSerTyGlu 533
OY 380 AATCCCTGGAGGTGACAAAGGCTGAGTACCGCTGCTACTCATCTTCTGTCAGT 439
Db 534 AlaAlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSer 553
OY 440 CAGCACTGTGACTCTCTTTATAAGTAAATAGTATAGTACACAGAGAAATTTCTCTTT 499
Db 554 GlnAsnPheAspMetPheValSerHisTySerSerValSerThrProProCysValHis 573
OY 500 CTTTACAAGCTATCAAGTCTCAAGATGACCAACTGTGCAAAACAAAGAAATTTGGGCC 559
Db 574 ValTyLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheThrPala 593
OY 560 ACCATTTGGATTGACGAGGTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTT 619
Db 594 SerMetMetGluAlaAlaSerCysProProAspTyValProProGluIlePheHisPhe 613
OY 620 GAAAGTACTGAGTGGATTACATTTGATGGATGCTCTACAGCTCATGATCTACAGCCT 679
Db 614 HisThrArgSerAspValArgLeuTyGlyMetIleTyLysProHisAlaLeuGlnPro 633
OY 680 GAAAGAAATATCTTACTGTCTGTTCATATATATGTTGGTCTCTCAGGTGAGTTGGTAA 739
Db 634 GlyLysLysAspProThrValLeuPheValTyGlyGlyProGlnValGlnLeuValAen 653
OY 740 RATCGTTTAAAGGATCAAGTATTCGGTTGAATACCTGATCCCTCTCTAGGTTATGTG 799
Db 654 AsnSerPheLysGlyIleLysTyLeuArgLeuAsnThrLeuAlaSerLeuGlyTyAla 673
OY 800 GTTGAGTATGATGACCAAGGGATCTGTCCAGGAGGCTTAAATTTGAAGCGCCTTT 859
Db 674 ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu 693
OY 860 AAATATATAAAGGCTCAATAGAAATGACATCAGTGAAGGACTCAATATCTAGCT 919
Db 694 LysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAla 713
OY 920 TCTGATATGATTTTATTGACTTAGATCTGTGGGATCCAGGCTGCTCTATGAGGA 979
Db 714 GluLysTyGlyPheIleAspLeuSerArgValAlaIleHisGlyTyPheSerTyGlyGly 733
OY 980 TACTCTCCCTGATGACCAATAGTAAAGTACAGAGTCAATCTTTCAGGTTGCTATTGCTGG 1039
Db 734 PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheCysValAlaIleAlaGly 753
OY 1040 GCGCCAGTCACTGTGGATCTTCTATGATACAGATACAG 1081
Db 754 AlaProValThrValTrpMetAlaTyAspThrGlyTyThr 767
```

## RESULT 6

```
US-09-976-674-23
; Sequence 23, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 492
; TYPE: PRT
```

```
i ORGANISM: Homo sapiens
US-09-976-674-23
Alignment Scores:
Pred. No.: 1,76e-135 Length: 892
Score: 1220.50 Matches: 236
Percent Similarity: 54.3% Conservative: 54
Best Local Similarity: 44.2% Mismatches: 69
Query Match: 63.0% Indels: 175
DB: 2 Gaps: 2
US-10-825-632-8 (1-1083) x US-09-976-674-23 (1-892)
OY 5 GAAGATGCCAGATCAGCTGAGTTCGCTACCTTTGTTCTTCCAAAGAAATTTGATAGATAT 64
Db 263 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 282
OY 65 TCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCACTGGTGGT---AAAATCTTT 121
Db 283 ThrGlyTyTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 302
OY 122 AGAATTCATATCAAGAAATGATCAATCTGAGTGGAAATTTATTCATGTTTACATCCCT 181
Db 303 ArgIleLeuTyGluGluValAspGluSerGluValGluValIleHisValProSerPro 322
OY 182 ATGTTGGAACAAGGAGGCGAGTTCATTCCTGTTATCTCTAAACAGGTACAGCAATCTCT 241
Db 323 AlaLeuGluGluArgLysThrAspSerTyArgTyTrpArgTyTrpArgTyTrpGlySerLysAsnPro 342
OY 242 AAGTCACTTTTAAAGTATCGAAATATGATTCATGCTGCTGAAAGGAGTACATA----- 295
Db 343 LysIleAlaLeuLysIleAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 362
OY 295 ----- 295
Db 363 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyIleAla 382
OY 295 ----- 295
Db 383 ArgAlaGlyTyTrpThrArgAspGlyLysTyAlaTrpAlaMetPheLeuAspArgProGln 402
OY 295 ----- 295
Db 403 GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu 422
OY 295 ----- 295
Db 423 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyValValTy 442
OY 295 ----- 295
Db 443 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyTrpPheProGlnSer 462
OY 295 ----- 295
Db 463 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 482
OY 295 ----- 295
Db 483 LeuTyLysValThrAlaValLeuLysSerGlnGlyTyArgTrpSerGluProPheSer 502
OY 295 ----- 295
Db 503 ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu 522
OY 296 ----- 319
Db 523 TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluThrLysLeuVal 542
OY 320 TATTTCAGGACCAAGGCTCCCTTTAGAGCATCAGCTGTACGTAGTCAGTTAGTA 379
Db 543 TyPheGlnGlyThrLysAspThrProLeuGluHisLeuTyValValSerTyGlu 562
```

QY 380 AATCTGGAGAGTGACAGGCTGACGTACCGTGGCTACTCAGTCTTCTGCTGATCAGT 439  
Db 563 AlaAlaGlyLeuIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSer 582  
QY 440 CAGCACTGTGACTCTTTTATAAGTAACTAGTAGTAACACAGAGAATCCACACTGTGTGCC 499  
Db 583 GlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrProProCysValHis 602  
QY 500 CTTTACAGCTATCAAGTCTGGAAGATGACCAACTTGCAAAACAAGAAATTTTGGGCC 559  
Db 603 ValTyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAla 622  
QY 560 ACCATTTTGGATTACAGAGGTCCTCTTCTGACTATACCTCTCCAGAAATTTTCTTTT 619  
Db 623 SerMetMetGluAlaAlaSerCysProProAspTyrValProProGluIlePheHisPhe 642  
QY 620 GAAATGACTACTGAGTTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCT 679  
Db 643 HisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnPro 662  
QY 680 GGAAGAATAATCTCTACTGCTGCTTCATATATATGTTGCTCCTCAGGTGCAGTTGGTGAAT 739  
Db 663 GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn 682  
QY 740 AATCGGTTTAAAGGAGTCAAGTATTTCGGCTTGAATACCTAGCCTCTCTAGGTTATGTG 799  
Db 683 AsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla 702  
QY 800 GTTGTAGTAGATACACACAGGGGATCCTGTGTCACGAGGGCTTAAATTTGAAGCGCTTT 859  
Db 703 ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu 722  
QY 860 AAATATAAATGGTCAATAGAAATTCAGCATCAGTGAAGGACTCCATATCTAGCT 919  
Db 723 LysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAla 742  
QY 920 TCTCGATATGATTTCATTCAGTATAGATCGTGTGGGCATCCACGGCTGTCTATGGAGGA 979  
Db 743 GluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGly 762  
QY 980 TACTCTCCTGATGCTGATTAATGACAGAGGTACAGATATCTTCAGGGTGTCTATGCTGGG 1039  
Db 763 PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGly 782  
QY 1040 GCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAG 1081  
Db 783 AlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 796

RESULT 7

US-09-976-674-27  
; Sequence 27, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 27  
; LENGTH: 892  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-27

Alignment Scores: 1.76e-135 Length: 892  
Pred. No.:

Score: 1220.50 Matches: 236  
Percent Similarity: 54.3% Conservative: 54  
Best Local Similarity: 44.2% Mismatches: 69  
Query Match: 63.0% Indels: 175  
DB: 2 Gaps: 2  
US-10-825-632-8 (1-1083) x US-09-976-674-27 (1-892)  
QY 5 GAAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCAGAGAATTTGATAGATAT 64  
Db 263 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 282  
QY 65 TCTGGCTATTGCTGTGTCCAAAAGCTCAAACTCCACAGTGGTGT---AAAATCTT 121  
Db 283 ThrGlyTyrTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuThrLeu 302  
QY 122 AGAATCTATATGAAGAAATGATGATCTGAGTGGAAATTTATCTGTTACATCCCT 181  
Db 303 ArgIleLeuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro 322  
QY 182 ATGTTGGAACAAGAGGCGCAGATTCTCCGTTTATCTTAAACACAGGTACAGCAATCT 241  
Db 323 AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 342  
QY 242 AAGTCACTTTTAAAGATGTCAAAATATGATGATCTGAAGGAGGATCATA----- 295  
Db 343 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 362  
QY 295 ----- 295  
Db 363 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 382  
QY 295 ----- 295  
Db 383 ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 402  
QY 295 ----- 295  
Db 403 GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu 422  
QY 295 ----- 295  
Db 423 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr 442  
QY 295 ----- 295  
Db 443 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 462  
QY 295 ----- 295  
Db 463 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 482  
QY 295 ----- 295  
Db 483 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 502  
QY 295 ----- 295  
Db 503 ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu 522  
QY 296 -----GTTGATGAAGTCAGAGGCTGTA 319  
Db 523 TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuVal 542  
QY 320 TATTTGAAGCCACCAAGACTCCCTTTAGACATCAGTCTAGTACAGTTACGTA 379  
Db 543 TyrPheGlnGlyThrLysAspTrpProLeuGluHisLeuTyrValValSerTyrGlu 562  
QY 380 AATCTGGAGAGGTGACAGGCTGACTGACCGGGCTACTCACATTTCTGCTGCATCAGT 439  
Db 563 AlaAlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSer 582  
QY 440 CAGCACTGTGACTCTTTTATAAGTAACTAGTAGTAACACAGAGAATCCACACTGTGTGCC 499



```
Db 583 GlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrProCysValHis 602
Qy 500 CTTTACAGCTATCAAGTCTGAGATGACCACTGCTCAAAACAAAGGAATTTGGGCC 559
Db 603 ValTyrLysLeuSerGlyProAspAspAspProLeuHisLysGlnProArgPheTrpAla 632
Qy 560 ACCATTGTCATTCACAGGTCCTCTTCCTGACTATCTCTCCGAAATTTCTCTTTT 619
Db 623 SerMetMetGluAlaAlaSerCysProProAspTyrValProProGluLeuPheHisPhe 642
Qy 620 GAAAGTACTACTGGATTACATTGTTGGGATGCTCTCAAGCCCTCATGATCTACAGCCT 679
Db 643 HisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnPro 662
Qy 680 GGAAGAAATATCTACTGCTGCTTCAATATATGTTGCTCTCAGGTGCAAGTGGTGAAT 739
Db 663 GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn 682
Qy 740 AATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTG 799
Db 683 AsnSerPheLysGlyLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla 702
Qy 800 GTTGTAGTATGATACACACAGGGATCCTGTCACCGAGGCTTAAATTTGAAGCGCTTT 859
Db 703 ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu 722
Qy 860 AATATAAATGGTCAATAGAAATTCACATCAGGTGGAAGGACTCCAAATATCTAGCT 919
Db 723 LysAsnGlnMetGlyGlnValGluLeuGluAspGlnValGluGlyLeuPheValAla 742
Qy 920 TCTCGATATGATTTCATTGACTTATGATCGTGGGATCCACCGCTGCTCAGGTGCTATCTGG 1039
Db 743 GluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGly 762
Qy 980 TACTCTCTCCCTGATGGCAATTAATCAGAGTTCAGATATCTCAGGCTGCTATCTGG 1039
Db 763 PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGly 782
Qy 1040 GCCCAGTCACTCTGTGATCTTCTATGATACAGGATACAGC 1081
Db 783 AlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 796

RESULT 8
; Sequence 11, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-976-674-11

Alignment Scores:
Pred. No.: 3,53e-134 Length: 661
Score: 1209.00 Matches: 244
Percent Similarity: 58.2% Conservative: 0
Best Local Similarity: 58.2% Mismatches: 0
Query Match: 62.4% Indels: 175
DB: 2 Gaps: 2
```

```
US-10-825-632-8 (1-1083) x US-09-976-674-11 (1-661)
Qy 2 GAAGAAGATGCCAGATCAGCTGGAGTCGGTACCTTTGTTCTCAAGAAGAATTTGATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
Qy 62 TATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCACAGTGGTGGTAAATTCCTT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283
Qy 122 AGAATTCATATGAAGAAATGATGATCTGAGGTGGAAATTTATTTCATCTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303
Qy 182 ATGTTGCAACAAGAGGAGGACAGATTCATTCCTTATCTTAAACAGGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
Qy 242 AAGTCTCACTTTAAGATGTCAGAAATATGATGATGCTGAAGGAGGATCATAT 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleAspVal 343
Qy 295 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
Qy 295 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
Qy 295 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
Qy 295 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
Qy 295 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
Qy 295 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
Qy 295 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
Qy 295 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
Qy 296 296
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
Qy 323 TTTGAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy 383 CTGAGAGGTGACAAAGCTGACTACCGTGGCTACTCATTTCTTGCTGCATCAGTCAG 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563
Qy 443 CACTGTGACTCTTTATAGTATAGTAACTAGTAAACAGAGAATCCACACTGTGTGCCCTT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 503 TACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCACAAACAAAGAAATTTGGGCCACC 562
```



```
QY 1019 TTCAGGTTCTCTATCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATAC 1078
Db 763 PheLysValAlaIleAlaGlyAlaProValThrValThrMetAlaTyrAspThrGlyTyr 782
QY 1079 ACG 1081
Db 783 Thr 783

RESULT 10
US-09-976-674-35
; Sequence 35, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junten, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-35

Alignment Scores:
Pred. No.: 3,61e-133 Length: 879
Score: 1201.00 Matches: 234
Percent Similarity: 55.1% Conservative: 53
Best Local Similarity: 44.9% Mismatches: 72
Query Match: 62.0% Indels: 162
D8: 2 Gaps: 3

US-10-825-632-8 (1-1083) x US-09-976-674-35 (1-879)
QY 5 GAAGATCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAAAGAAATTTGATAGATAT 64
Db 263 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluGluPheAspArgPhe 282
QY 65 TCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGT---AAAATTCCT 121
Db 283 ThrGlyTyrTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 302
QY 122 AGAATTCATATGAAGAAATGATGATCTGAGTGGAAATTTATTCATGATACATCCCT 181
Db 303 ArgIleLeuTyrGluGluValAlaAspGluSerGluValIleHisValProSerPro 322
QY 182 ATGTTGGAACAGAGGAGGCAGATTCATTCCGTTATCCTAAACAGGTACAGCAATCCT 241
Db 323 AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 342
QY 242 AAAGTCACTTTAAGATGTCAGAAATATGATTCATGCTGAGGAGGATCATAGTTGAT 301
Db 343 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 362
QY 302 GAAGTCAGAGGCTG----- 316
Db 363 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 382
QY 316 ----- 316
Db 383 ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 402
QY 316 ----- 316
Db 403 GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu 422
```

```
QY 317 -----CTATAT 322
Db 423 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr 442
QY 323 TTTGAA----- 328
Db 443 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 462
QY 328 ----- 328
Db 463 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 482
QY 328 ----- 328
Db 483 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 502
QY 328 ----- 328
Db 503 ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu 522
QY 329 -----GGCCCAAAGACTCCCTTTAGAGCATCAC 358
Db 523 TrpGluValLeuAlaArgHisGlySerLysGlyThrLysAspThrProLeuGluHisHis 542
QY 359 CTGTAGTAGTCAGTACGTAATCCTCGAGAGGTGACAAGGCTGACGACCGTGGCTAC 418
Db 543 LeuTyrValValSerTyrGluAlaGlyGluIleValArgLeuThrThrProGlyPhe 562
QY 419 TCACATTTCTGTCATCAGTCAGCAGCTGTGACTCTTTTATAAGTAAAGTATAGTAAACAG 478
Db 563 SerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal 582
QY 479 AAGAATCCACTGTGTGTCCTTTACAAGCTATCAAGTCTCGAAGATGACCCCACTTGC 538
Db 583 SerThrProProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHis 602
QY 539 AAAACAAAGGAATTTGGGCCACCATTTGGATTTCAGCAGGCTCTCTCTCCTACATATCT 598
Db 603 LysGlnProArgPheTrpAlaSerMetGluAlaAlaSerCysProAspTyrVal 622
QY 599 CCTCAGAAATTTCTCTTTGAAAGTACTACTGGATTTCATTTGATGGGATGCTCTAC 658
Db 623 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyr 642
QY 659 AAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGCTGCTGTTCATATATCGTGT 718
Db 643 LysProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGly 662
QY 719 CCTCAGGTGAGTGGTGAATAATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 778
Db 663 ProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThr 682
QY 779 CTAGCCTCTCTAGGTTATGTTGGTTGTAGTAGACACACGCGGATCCTGTCA CGAGGG 838
Db 683 LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly 702
QY 839 CTTAAATTTGAAGCGCCCTTTAAATATAAATGCGTCAAAATAGAAATTCAGCATCAGGTG 898
Db 703 LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal 722
QY 899 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGATCTTAGATCGTGGGCATC 958
Db 723 GluGlyLeuPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle 742
QY 959 CAGGCTGTCTTATGAGGATACCTCTCCCTGATGGCATTAAATGACAGATCATGATATC 1018
Db 743 HisGlyTrpSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal 762
QY 1019 TTCAGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTTATGATACAGGATAC 1078
Db 763 PheLysValAlaIleAlaGlyAlaProValThrValThrMetAlaTyrAspThrGlyTyr 782
```

Alignment Scores:	4.74e-125	Length:	832
Pred. No.:			



```

US-10-825-632-8 (1-1083) x US-09-976-674-31 (1-832)
QY 5 GAAGATGCCAGATCAGCTGGAGTCGCTTCTCCCAAGAAATTTGATAGATAT 64
Db 263 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 282
QY 65 TCTGGCTATTGCTGGTCCAAAAGCTGAAACAACTCCCACTGGTGGT---AAAATCTTT 121
Db 283 ThrGlyTyrTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuYsThrLeu 302
QY 122 AGAATCTATATCAAGAAATGATCAATCTGAGTGGAAATATTTCATGTTACATCCCT 181
Db 303 ArgIleLeuTyrGluGluValAspGluSerGluValGluValIleHisValProSerPro 322
QY 182 ATGTTGGAACAGGAGGCGACATTCATTCCGTTATCTTAAACAGATCAGCAAAATCTT 241
Db 323 AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 342
QY 242 AAAGTCACTTTTAAAGATGTCAGAAATTAATGATGCTGAAGAGGATCATA----- 295
Db 343 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 362
QY 295 ----- 295
Db 363 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 382
QY 295 ----- 295
Db 383 ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 402
QY 295 ----- 295
Db 403 GlnTrpLeuGlnLeuValLeuProProAlaLeuPheIleProSerThrGluAsnGlu 422
QY 295 ----- 295
Db 423 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr 442
QY 295 ----- 295
Db 443 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 462
QY 295 ----- 295
Db 463 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 482
QY 295 ----- 295
Db 483 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 502
QY 295 ----- 295
Db 503 ProGlyGluAspGluPheLysCysProIleLysGluGluLeuAlaLeuThrSerGlyGlu 522
QY 296 -----GTTGATGAAGTCAGAGGCTGGTA 319
Db 523 TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuVal 542
QY 320 TATTTTGAAGGACCAAGATCCCTTTAGAGCATCCTGTAGCTAGTACGTACGTACGTA 379
Db 543 TyrPheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGlu 562
QY 380 AATCTGGAGGTGACAGGCTGACTGACCGTCTACTCACATTTCTGTCGTCATCAGT 439
Db 563 AlaAlaGlyLeuIleValArgLeuThrProGlyPheSerHisSerCysSerMetSer 582
QY 440 CAGCACTGCTACTTCTTTAAGTATAGTAAACAGAAATCCACACTGTGTCTCC 499
Db 583 GlnAsnPheAspMetPheValSerHisTyrSerValSerValSerThrProCysValHis 602
QY 500 CTTTACAGCTATCAAGTCTGGAATGACCCCACTTCGAAACAAAGGAATTTGGGCC 559

```

```

RESULT 14
US-09-976-674-37
; Sequence 37, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 819
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-37

Alignment Scores:
Pred. No.: 9,69e-123 Length: 819
Score: 1113.00 Matches: 220
Percent Similarity: 52.9% Conservative: 51
Best Local Similarity: 43.0% Mismatches: 79
Query Match: 57.4% Indels: 162
DB: 2 Gaps: 3

US-10-825-632-8 (1-1083) x US-09-976-674-37 (1-819)
QY 5 GAAGATGCCAGATCAGCTGGAGTCGCTTCTCCCAAGAAATTTGATAGATAT 64

```

Db 263 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 282  
QY 65 TCTGCTATTGGTGGTCTCCAAAGCTGAACAATCCAGTGGTGGT---AAATTTCTT 121  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 121  
283 ThrGlyTyrTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 302  
QY 122 AGATTCATATCAAGAAATGATGAATCTGAGGTGGAAATATTTCATGTTTACATCCCT 181  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181  
303 ArgLeuLysGluGluValAspGluSerGluValGluValIleHisValProSerPro 322  
QY 182 ATGTTGCAACAGGAGGCGAGATTCTTCCTGTTATCTTAAACAGGTACAGCAATCTCT 241  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 241  
323 AlaLeuGluGluLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 342  
QY 242 AAATGCTCTTTAAGATGTCAGAAATAATGATGCTGAAGGAGGATCATAGTTGAT 301  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 301  
343 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 362  
QY 302 GAAGTCAGAGGCTG--- 316  
Db :||:|:| 316  
363 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 382  
QY 316 --- 316  
383 ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 402  
QY 316 --- 316  
403 GlnTrpLeuGlnLeuValLeuLeuProProAlaLeuPheIleProSerThrGluAsnGlu 422  
QY 317 -----GTATAT 322  
423 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTrp 442  
QY 323 TTTGAA--- 328  
443 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 462  
QY 328 ----- 328  
463 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 482  
QY 328 ----- 328  
483 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 502  
QY 328 ----- 328  
503 ProGlyGluAspGluPheLysCysProLysGluGluIleAlaLeuThrSerGlyGlu 522  
QY 329 -----GCCACCAAGACTCCCTTTAGAGCATCAC 358  
523 TrpGluValLeuAlaArgHisGlySerLysGlyThrLysAspThrProLeuGluHisHis 542  
QY 359 CTGTACGTAGTCAGTTAGCTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTAC 418  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 418  
543 LeuTyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPhe 562  
QY 419 TCACATCTTGCTGCATCAGTCAGTCAGTCTGACTTCTTTATAGTAAAGTATAGTACAC 478  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 478  
563 SerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal 582  
QY 479 AAGATCCACACTGTGTCTCCCTTTACAAAGTATCAAGTCTCAAGATCACCAACTGTC 538  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 538  
583 SerThrProProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHis 602  
QY 539 AAAACAAAGGAATTTGGCCACCATTTTGGATTTCAGCAGGCTCTCTTCGACTACTACT 598  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 598  
603 LysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysAspProAspTyrVal 622  
QY 599 CTTCCAGAAATTTCTCTTTTGAAGTACTACTGAGTATTTACATTTGATGGGATGCTCTAC 658  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 658  
623 ProProGluIlePheHisPheHisThrArgSerAspValArgLysTyrGlyMetIleTyr 642

QY 659 AAGCCTCATGATCTACAGCTGTGAAGAAATATCTCTGCTGCTGTTTCATATATGGTGGT 718  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 718  
643 LysProHisAlaLeuGlnProGlyLysHisProThrValLeuPheValTyrGlyGly 662  
QY 719 CTTGAGTGCAGTTGGTGAATTAATCGTTTAAAGGAGTCAAGTATTTCCGCTTCAATACC 778  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 778  
663 ProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLysAsnThr 682  
QY 779 CTAGCCTCTCTAGTTTATGTTGTTAGTGTATAGTACACACAGGGATCTGTCCACCGAGG 838  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 838  
683 LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly 702  
QY 839 CTTAAATTTGAAGGCGCCTTTAAATATAAATGGTCAATAGAAATTCAGCATCAGGTG 898  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 898  
703 LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal 722  
QY 899 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTTGATCTGTGGGATC 958  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 958  
723 GluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle 742  
QY 959 CACGGCTGCTCTATGGAGGATACCTCTCCCTGATGCGATTATGCAGAGGTGAGATATC 1018  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1018  
743 HisGlyTyrSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal 762  
QY 1019 TTTGAGGCTTGTCTATTGCTGGGCGCCAGTCACCTCTG 1054  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1054  
763 PheLysAlaGlnProLeuAlaTyrProProArgLeu 774

RESULT 15  
US-09-976-674-39  
; Sequence 39, Application US/09976674  
; Patent No. 6844180  
; GENERAL INFORMATION:  
; APPLICANT: Oi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 39  
; LENGTH: 819  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-39

Alignment Scores:  
Pred. No.: 9,69e-123 Length: 819  
Score: 1113.00 Matches: 220  
Percent Similarity: 52.9% Conservative: 51  
Best Local Similarity: 43.0% Mismatches: 79  
Query Match: 57.4% Indels: 162  
DB: 2 Gaps: 3

US-10-825-632-8 (1-1083) x US-09-976-674-39 (1-819)

QY 5 GAAGATGCCAGATCAGCTGGAGTGCCTACCTTCTCCCAAGAAATTTGATATAT 64  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 64  
263 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 282  
QY 65 TCTGCTATTGGTGGTGTGCCAAAGCTGAAACAACTCCAGTGGTGGT---AAATTTCTT 121  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 121  
283 ThrGlyTyrTrpTyrProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeu 302  
QY 122 AGAATTCATATGAAGAAATGATGATCTGAGGTGGAATTTTCATCTTACATCCCT 181  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181  
303 ArgIleLeuTyrGluGluValAspGluSerGluValGluIleHisValProSerPro 322



```
QY 182 ATGTTGAAACAAAGGAGGCGAGATTTCATTCGTTATCTCTAAACAGGTACAGCAATCCT 241
Db 323 AlaLeuGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 342
QY 242 AAGTCACCTTTTAAAGATCTCAGAAATATGATGATCTCAAGAGAGGATCATAGTTGAT 301
Db 343 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 362
QY 302 GAAGTCAGAGAGGCTG----- 316
Db 363 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 382
QY 316 ----- 316
Db 383 ArgAlaGlyTrpThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 402
QY 316 ----- 316
Db 403 GlnTrpLeuGlnLeuValLeuProProAlaLeuPheIleProSerThrGluAsnGlu 422
QY 317 -----GTATAT 322
Db 423 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr 442
QY 323 TTGGA----- 328
Db 443 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 462
QY 328 ----- 328
Db 463 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 482
QY 328 ----- 328
Db 483 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 502
QY 328 ----- 328
Db 503 ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu 522
QY 329 -----CGCACCAAGACTCCCTTTAGAGCATCAC 358
Db 523 TrpGluValLeuAlaArgHisGlySerLysGlyThrLysAspThrProLeuGluHisHis 542
QY 359 CTGTACGTAGTCAGTTTACCTAAATCCTCGAGAGGTGACCAAGCTGACTGACCGTGGCTAC 418
Db 543 LeuTyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThrProGlyPhe 562
QY 419 TCACATCTTCTGTCATCAGTCAGCAGTCGTGACTCTTTTATAGTAAGTATAGTAACCAG 478
Db 563 SerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerVal 582
QY 479 AAGAATCCACACTGTGTGCTCTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGC 538
Db 583 SerThrProCysValHisValTyrLysLeuSerGlyProAspAspProLeuHis 602
QY 539 AAAAAAGAAATTTGGGCCACCAATTTGGATTACAGCAGTCCTCTTCTGACTATACT 598
Db 603 LysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProProAspTyrVal 622
QY 599 CTTCCAGAAATTTCTCTTCAAGTACTACTGATTTCATTGTATGGATGCTCTAC 658
Db 623 ProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyr 642
QY 659 AAGCCTCATGATCTCAGCCTCGAAAGAAATATCCTACTGTGCTGTTTCATATATGGTGGT 718
Db 643 LysProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyGly 662
QY 719 CCTCAGTGTGAGTAAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 778
Db 663 ProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThr 682
```

Search completed: May 2, 2006, 01:42:33  
Job time : 58.1838 secs

```
QY 779 CTAGCCTCTAGGTTATGTGTTAGTATAGACACAGGGGATCCTGTCCCGAGGG 838
Db 683 LeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGly 702
QY 839 CTTAAATTTGAAGGGCCCTTTAAATATATAAATGGGTCAAAATAGAAATTCAGCATCAGGTG 898
Db 703 LeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGluAspGlnVal 722
QY 899 GAAGGACTCAATATCTAGCTTCTCGATATGATTTTCATTGACTTATGATCGTGTGGGCATC 958
Db 723 GluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArgValAlaIle 742
QY 959 CACGGCTGTGCTTATGGAGATACCTCTCCCTGTAGTGGCATTAAATCGACAGGTGCAGATATC 1018
Db 743 HisGlyTrpSerTyrGlyGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnVal 762
QY 1019 TTCAGGTTGCTATTGCTGGGCCCCAGTCACCTCTG 1054
Db 763 PheLysAlaGlnProLeuAlaTyrProProArgLeu 774
```

**THIS PAGE BLANK (USPTO)**

Sequence 2, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 620, App  
Sequence 7, Appli  
Sequence 3, Appli  
Sequence 619, App  
Sequence 3, Appli  
Sequence 23, Appli  
Sequence 27, Appli  
Sequence 23, Appli  
Sequence 27, Appli  
Sequence 12, Appli  
Sequence 2, Appli  
Sequence 22, App  
Sequence 226, App  
Sequence 11, Appli  
Sequence 11, Appli  
Sequence 4, Appli  
Sequence 33, Appli  
Sequence 35, Appli  
Sequence 33, Appli  
Sequence 35, Appli  
Sequence 19, Appli  
Sequence 19, Appli  
Sequence 29, Appli  
Sequence 31, Appli  
Sequence 29, Appli  
Sequence 31, Appli  
Sequence 37, Appli  
Sequence 39, Appli  
Sequence 37, Appli  
Sequence 39, Appli  
Sequence 4, Appli

580 6 US-11-140-224-2  
690 3 US-09-976-674-7  
690 5 US-10-982-512-7  
689 4 US-10-072-012-620  
689 4 US-10-415-122-7  
863 3 US-09-976-674-3  
863 4 US-10-072-012-619  
863 5 US-10-982-512-3  
892 3 US-09-976-674-23  
892 3 US-09-976-674-27  
892 5 US-10-982-512-23  
892 5 US-10-982-512-27  
892 5 US-10-433-757-12  
892 4 US-10-072-012-224  
863 4 US-10-072-012-226  
863 3 US-09-976-674-11  
661 5 US-10-982-512-11  
661 5 US-10-415-122-4  
869 4 US-09-976-674-33  
879 3 US-09-976-674-35  
879 5 US-10-982-512-33  
879 5 US-10-982-512-35  
879 5 US-09-976-674-19  
658 5 US-10-982-512-19  
658 3 US-09-976-674-29  
832 3 US-09-976-674-31  
832 5 US-10-982-512-29  
832 5 US-10-982-512-31  
819 3 US-09-976-674-37  
819 3 US-09-976-674-39  
819 5 US-10-982-512-37  
819 5 US-10-982-512-39  
310 3 US-09-993-959-4

12 1495 77.1  
13 1226 63.3  
14 1226 63.3  
15 1220.5 63.0  
16 1220.5 63.0  
17 1220.5 63.0  
18 1220.5 63.0  
19 1220.5 63.0  
20 1220.5 63.0  
21 1220.5 63.0  
22 1220.5 63.0  
23 1220.5 63.0  
24 1220.5 63.0  
25 1212.5 62.6  
26 1212.5 62.6  
27 1212.5 62.6  
28 1209 62.4  
29 1209 62.4  
30 1204.5 62.2  
31 1201 62.0  
32 1201 62.0  
33 1201 62.0  
34 1201 62.0  
35 1199.5 61.9  
36 1199.5 61.9  
37 1132.5 58.4  
38 1132.5 58.4  
39 1132.5 58.4  
40 1132.5 58.4  
41 1113 57.4  
42 1113 57.4  
43 1113 57.4  
44 1113 57.4  
45 1086.5 56.1

ALIGNMENTS

RESULT 1  
US-10-825-632-7  
; Sequence 7, Application US/10825632  
; Publication No. US20040191826A1  
; GENERAL INFORMATION:  
; APPLICANT: ABBOTT, Catherine Anne  
; APPLICANT: CORRELL, Mark Douglas  
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES  
; FILE REFERENCE: FCSB-100-Div. 1  
; CURRENT APPLICATION NUMBER: US/10/825,632  
; CURRENT FILING DATE: 2004-04-15  
; PRIOR APPLICATION NUMBER: US 10/070,464  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: PCT/AU00/01085  
; PRIOR FILING DATE: 2000-09-11  
; PRIOR APPLICATION NUMBER: AU PQ5709  
; PRIOR FILING DATE: 2000-02-18  
; PRIOR APPLICATION NUMBER: AU PQ2762  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-825-632-7

Alignment Scores:  
Pred. No.: 6,53e-192 Length: 360  
Score: 1933.00 Matches: 360  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 99.7% Indels: 0  
DB: 4 Gaps: 0

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 02:43:09 ; Search time 31.0173 Seconds  
(without alignments)  
2917.786 Million cell updates/sec

Title: US-10-825-632-8  
Perfect score: 1938  
Sequence: 1 ggaagaagatgcagatcag.....tatgatacaggatacacgga 1083

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_n2p\_model -DEV=xlp  
-Q=/abs/ABSWEB\_spool/US10825632/runat\_01052006\_105957\_3449/app\_query.fasta\_1  
-DB=Published Applications AA Main -OPMT=fastan -SUFFIX=n2p.rapbm  
-MINMATCH=0.1 -LOOPEL=0 -DOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs602p  
-USER=US10825632 @CGN 1 1 503 @runat 01052006 105957\_3449 -NCPU=6 -ICPU=3  
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1933	99.7	360	4	US-10-825-632-7
2	1836.5	94.8	882	3	US-09-976-674-1
3	1836.5	94.8	882	4	US-10-054-776-2
4	1836.5	94.8	882	4	US-10-170-789-38
5	1836.5	94.8	882	4	US-10-311-035-9
6	1836.5	94.8	882	4	US-10-072-012-622
7	1836.5	94.8	882	4	US-10-415-122-6
8	1836.5	94.8	882	4	US-10-825-632-1
9	1836.5	94.8	882	5	US-10-982-512-1
10	1756.5	90.6	883	4	US-10-072-012-621
11	1495	77.1	580	4	US-10-275-505-2

US-10-825-632-8 (1-1083) x US-10-825-632-7 (1-360)

```
QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGA 61
Db 1 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 20
QY 62 TATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACCTCCAGCTGGTGTAAATTCCT 121
Db 21 Ty-SerGlyTyrTrpTrpCysProLysAlaGluThrProSerGlyGlyLysLeu 40
QY 122 AGAATTCATATGAAGAAATCATGAATCTGAGTGGGAATTTATTCATGTTACATCCCT 181
Db 41 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 60
QY 182 ATGTTGGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
Db 61 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 80
QY 242 AAGTCACATTTTAAGATCTCAGAAATATGATGCTGAAGAAAGGATCATAGTTGAT 301
Db 81 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleValAsp 100
QY 302 GAAGTCAGAGGCTGGTATATTTGAAGGCCAACAAAGACTCCCTTTAGAGCATCACCTG 361
Db 101 GluValArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeu 120
QY 362 TAGCTAGTCAGTTAGTAAATCCTGGAGAGTGACAGGCTGACTGACCGTGGCTACTCA 421
Db 121 TyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSer 140
QY 422 CATCTTCTGCTGATCAGTCAGCACTGTGACTCTCTTTATAGTAAATAGTATAGTAACAGAG 481
Db 141 HisSerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLys 160
QY 482 AATCACACTGTGTGCTCTTACAAGCTATCAAGTCTCAAGTCTCAAGTCAAGTCAAGTCAAG 541
Db 161 AsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 180
QY 542 ACAAGGAAATTTGGGCCACCATTTTGATTCAGCAGGCTCCTCTTCCTGACTATACCTCT 601
Db 181 ThrLysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrPro 200
QY 602 CCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAG 661
Db 201 ProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLys 220
QY 662 CCTCATGATCTACAGCTCGAAAGAAATATCTCTACTGCTGCTTCTCATATATGCTGGTCT 721
Db 221 ProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyPro 240
QY 722 CAGGTGCAAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTA 781
Db 241 GlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeu 260
QY 782 GCCTCTCTAGTTATGCTGTTGTAGTATAGTACACAGGGATCTGTCACCGAGGCTT 841
Db 261 AlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeu 280
QY 842 AAATTTGAGGCCCTTTAAATATAAATGGTCAAAATAGAAATAGCAATCATCAGGTGGAA 901
Db 281 LysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGlu 300
QY 902 GGACTCCATATCTAGCTTCTCGATATGATTTTCAATTTGACTTAGATCGTGGGAGTCCAC 961
Db 301 GlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHis 320
QY 962 GGCTGCTCTATGGAGGATPACTCTCCCTGATGGCATTATGACAGAGGTGAGATATCTTC 1021
Db 321 GlyTrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePhe 340
QY 1022 AGGTTTGTCTATGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081
Db 341 ArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 360
```

RESULT 2  
US-09-976-674-1  
; Sequence 1, Application US/09976674  
; Patent No. US20020115843A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; FILE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; CURRENT APPLICATION NUMBER: US/09/976,674  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-976-674-1

Alignment Scores:  
Pred. No.: 9,51e-182 Length: 882  
Score: 1836.50 Matches: 360  
Percent Similarity: 67.5% Conservative: 0  
Best Local Similarity: 67.5% Mismatches: 0  
Query Match: 94.8% Indels: 173  
DB: 3 Gaps: 1

US-10-825-632-8 (1-1083) x US-09-976-674-1 (1-882)

```
QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACCTCCAGCTGGTGTAAATTCCT 121
Db 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283
QY 122 AGAATTCATATGAAGAAATGATGAATCTGAGTGGAAATATTATTCATCTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303
QY 182 ATGTTGGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACATTTTAAGATCTCAGAAATATGATGCTGAAGAAAGGATCATAGTTGAT 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleValAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
```

295 ----- 295  
444 HisGluGluGluGluGluPhePheAlaSerGluCysLeuThrGlyPheArgHisLeu 463  
295 ----- 295  
464 TyrIleThrSerIleLeuLeuGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483  
295 ----- 295  
484 AlaProSerAppPheLysCysProIleLysGluGluLeuAlaIleThrSerGlyGluTrp 503  
296 -----GTTGATCAAGTCAGAAAGCGTGGTATAT 322  
504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523  
323 TTGAGGACCAAGAGACTCCCTTTAGAGATCACCTGTACGTAGTACGTAAAT 382  
524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543  
383 CCTGAGAGGTGACAAAGCTGACTGACCGTGGTACTCACATTCCTGCTGCATCAGTCAG 442  
544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysSilleSerGln 563  
443 CACTGTGACTCTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCT 502  
564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
503 TACAAGCTATCAAGTCCTGAAGATGACCAACTTCGAAACAAAGGAATTTGGGCCACC 562  
584 TyrIleLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
563 ATTTTGATTCACAGTCTCTCTGACTATATCTCTCCAGAAATTTCTTTTGA 622  
604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
623 AGTACTACTGATTTACATTTGATGGATGCTCTACAGCTCATGATCAGCTGGA 682  
624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
683 AAGAAATATCTACTGCTGCTTCATATATGTTGGTCTCAGGTGAGTGTGTAATAT 742  
644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663  
743 CGCTTTAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTTATGCTGT 802  
664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
803 GTAGTGATAGACAAAGGGATCCTGTCCAGGGCTTAAATTTGAAGCGCCTTTAAA 862  
684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
863 TATAAATGGGTCAATAGAATTCACCATCAGGTGGAAGACTCCAAATATCTAGCTTCT 922  
704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
923 CGATATGATTTCAATGATTCATGATCGTGTGGGCTCCACCGCTGCTCTATGGAGATAC 982  
724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743  
983 CTCTCCCTGATGCTATTAATCAGAGGTTCAGATATCTTCAGGGTTCATGCTGGGGCC 1042  
744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
1043 CCAGTCACTCTGTGGATCTTCTATGATACAGATACAG 1081  
764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

RESULT 3  
US-10-054-776-2  
; Sequence 2, Application US/10054776  
; Publication No. US20030165818A1

GENERAL INFORMATION:  
; APPLICANT: Mark Robert Edbrooke  
; APPLICANT: Alan Peter Lewis  
; TITLE OF INVENTION: NOVEL PROTEIN  
; FILE REFERENCE: QG1042US  
; CURRENT APPLICATION NUMBER: US/10/054,776  
; CURRENT FILING DATE: 2002-01-23  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-054-776-2  
Alignment Scores: 9,51e-182 Length: 882  
Pred. No.: 1836,50 Matches: 360  
Score: 67,5% Conservative: 0  
Percent Similarity: 67,5% Mismatches: 0  
Best Local Similarity: 94,8% Indels: 173  
Query Match: 4 Gaps: 1  
DB:  
US-10-825-632-8 (1-1083) x US-10-054-776-2 (1-882)  
QY 2 GAAGAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGA 61  
Db 244 GluGluAspAlaArgSerAlaGlyValAlaIleThrPheValLeuGlnGluPheAspArg 263  
QY 62 TATTCTCGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTCCT 121  
Db 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysLeu 283  
QY 122 AGAATTCCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCT 181  
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303  
QY 182 ATGTTGGAACAAGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACAGCAATCTCT 241  
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323  
QY 242 AAGTCACTTTTAAGATGTGCAAAATATGATTCATGATGCTGCAAGGAAGGATCATA 295  
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343  
QY 295 ----- 295  
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363  
QY 295 ----- 295  
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383  
QY 295 ----- 295  
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403  
QY 295 ----- 295  
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423  
QY 295 ----- 295  
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443  
QY 295 ----- 295  
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463  
QY 295 ----- 295  
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483

```
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTyr 503
QY 296 ----- 296
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValAlaGluValTyr 523
QY 323 TTTGAAGGACCAAAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAGTTACGTAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisIleuTyValValSerTyValAsn 543
QY 383 CTGGAGAGGTCACAAAGGCTGACTGACCGTGGTACTCAATCTTCTGTCGACGTCAG 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysValSerGln 563
QY 443 CACTGTGACTCTTTTATAGTAGTATAGTACAGAGAAATCCACACTGTGTGCTCTT 502
Db 564 HisCysAspPhePheIleSerLysTySerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAAGCTATCAAGTCTGAGATGACCAACTTGCAAAACAAGGAATTTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspProThrCysLeuThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGATTCCAGAGTCCCTTCCTGACTATCTCCAGAAATTTTCTCTTTGAA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AAGAATATCTACTGCTGTTTCATATATGTTGGTCTCTCAGTGCAGTTCGTCGTAAT 742
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
QY 743 CGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTTGTT 802
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTATAGAACACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCTTAA 862
Db 684 ValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATAAATGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 923 CGATATGATTTCACTGACTTAGATCGTGTGGCATCCACGCTGCTCTATGAGGATAC 982
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
QY 983 CTCTCCCTGATGGCATTATGACAGGTTCAGATATCTTCAGGTTCTCTATTCGTTGGGCC 1042
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCAGTCACTGTGTGATCTTCTATGATACAGGATACAG 1081
Db 764 ProValThrLeuTyrPheThrAspThrGlyTyrThr 776

RESULT 4
US-10-170-789-38
; Sequence 38, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF

FILE REFERENCE: 10448-191001
CURRENT APPLICATION NUMBER: US/10/170,789
CURRENT FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: US 09/797,039
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/06525
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/186,061
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 09/882,166
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19269
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/212,078
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 09/934,406
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: PCT/US01/26052
PRIOR FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 60/226,740
PRIOR FILING DATE: 2000-08-21
PRIOR APPLICATION NUMBER: US 09/861,801
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: PCT/US01/16549
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 60/205,508
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/801,267
PRIOR FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: PCT/US01/07138
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: US 60/187,454
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/829,671
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: PCT/US01/40483
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: US 60/197,508
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: US 09/961,721
PRIOR FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: PCT/US01/29904
PRIOR FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 60/235,023
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US 10/045,367
PRIOR FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/246,561
PRIOR FILING DATE: 2000-11-07
PRIOR APPLICATION NUMBER: US 09/801,275
PRIOR FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: PCT/US01/07074
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: US 60/187,420
NUMBER OF SEQ ID NOS: 63
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 38
LENGTH: 882
TYPE: PRT
ORGANISM: Homo sapiens
US-10-170-789-38

Alignment Scores:
Pred. No.: 9,51e-182 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 4 Gaps: 1

US-10-825-632-8 (1-1083) x US-10-170-789-38 (1-882)
```

QY 2 GAAGAAGATCCAGATCAGTGGAGTCGCTACCTTGTCTCCAAAGAAATTTGATAGA 61  
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGluPheAspArg 263  
QY 62 TATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGTAAATCTT 121  
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283  
QY 122 AGAATCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCT 181  
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303  
QY 182 ATGTGGAAACAAGAGGCGAGATTCATTCCTTATCTCTAAACAGGTACAGCAATCTCT 241  
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323  
QY 242 AAAGTCACTTTTAAGATGTCAGAAATAATGATGCTGAGGAGGAGGATCAT- 295  
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleAspVal 343  
QY 295 ----- 295  
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363  
QY 295 ----- 295  
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383  
QY 295 ----- 295  
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403  
QY 295 ----- 295  
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423  
QY 295 ----- 295  
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443  
QY 295 ----- 295  
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463  
QY 295 ----- 295  
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483  
QY 295 ----- 295  
Db 484 AlaProSerAspPheLysCysProLysLysGluGluIleAlaIleThrSerGlyGluTrp 503  
QY 296 -----GTTGATGAAGTCAGAGGCTGGTATAT 322  
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgGlyLeuValTyr 523  
QY 323 TTGAGAGGCACCAAGACTCCCTTTAGAGCATCACCCTGAGTACGTAGTCAGTACGTAAT 382  
Db 524 PheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsn 543  
QY 383 CTGGAGAGGTCAAGGCTGACTGACCGTGGCTACTCACATCTTGTGTCAGTCAGTCAG 442  
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563  
QY 443 CACTGTGACTCTTTTATAAGTAGTAGTAACACAGAAATCCACACTGTGTGTCCTT 502  
Db 564 HisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
QY 503 TACAAGCTATCAAGTCTCAAGATGACCAACTCTCAAAACAAAGGAATTTTGGGCCACC 562  
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
QY 563 ATTTTGGATTTCAGAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAA 622

Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
QY 623 AGTACTACTGATTTACATTGTATGGGATCTCTACAGCCTCATGATCTACAGCCTGGA 682  
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
QY 683 AAGAAATATCTACTGTCTGTTTCATATATGGTGGTCTCAGGTGCAGTGGTGAATAAT 742  
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663  
QY 743 CGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGTATTATGTTT 802  
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
QY 803 GTAGTGATAGACACAGGGATCTGTCCAGGAGGCTTAAATTTCAAGCGCCTTTAAA 862  
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
QY 863 TATAAATGGTCAAAATAGAAATTCAGCATCAGGTGGAGGACTCCAATATCTAGCTTCT 922  
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
QY 923 CGATATGATTTTCATTGACTTAGATCGTGTGGGATCCACGGTGGTCTCTATGAGGATAC 982  
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743  
QY 983 CTCCTCCCTGATGGCATTAAATGCAGAGTTCAGATATCTTCAGGGTTCCTATGCTGGGCG 1042  
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
QY 1043 CCAGTCTCTGTGGATCTTCTATGATACAGGATACAG 1081  
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

## RESULT 5

US-10-311-035-9  
; Sequence 9, Application US/10311035  
; Publication No. US20040023243A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: YUE, Henry  
; APPLICANT: ELLIOTT, Vicki  
; APPLICANT: GANDHI, Ameena R.  
; APPLICANT: LAL, Preeti  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: DELEGANE, Angelo M.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: NGUYEN, Dannel B.  
; APPLICANT: LEE, Ernestine A.  
; APPLICANT: HAFALIA, April  
; APPLICANT: KHAN, Farrah A.  
; APPLICANT: CHAWLA, Narinder K.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: LU, Dyuong Aina M.  
; APPLICANT: ARVIZU, Chandra S.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: WALSH, Roderick T.  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: LU, Yan  
; APPLICANT: RAMKUMAR, Jayalaximi  
; APPLICANT: XU Yuming  
; APPLICANT: REDDY, Roopa  
; APPLICANT: DAS, Depopriya  
; APPLICANT: KEARNEY, Liam  
; APPLICANT: KALLICK, Deborah A.  
; TITLE OF INVENTION: Proteases  
; FILE REFERENCE: PI-0123 PCT  
; CURRENT APPLICATION NUMBER: US/10/311,035  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946  
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14



```

; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1
US-10-311-035-9

```

```

Alignment Scores:
Pred. No.: 9,51e-182 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 4 Gaps: 1

```

US-10-825-632-8 (1-1083) x US-10-311-035-9 (1-882)

```

QY 2 GAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAATTGTATAGA 61
Db 244 GUGUUSPALAARGSERALAGLYVALAAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCCAAGTGGTGTGTAATAATTC 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrProSerGlyGlyLysLeu 283
QY 122 AGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTTCATGTTACATCCCT 181
Db 284 ArgIleLeuTyrGluLysAspGluSerGluValGluIleHisValThrSerPro 303
QY 182 ATGTTGGAAACAGGAGGAGATTCATTCGTTATCCTAAACAGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCACCTTTAAGATGTCAGAAATAATGATGATGCTGAAGAGAGATCATA----- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
QY 295 ----- 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503

```

```

QY 296 -----GTTGATGAAGTCCAGAGGCTGCTATAT 322
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
QY 323 TTGAAGGCACCAAGACTCCCTTTAGAGCATCACCCTGCTAGCTAGTACGTTACGTAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
QY 383 CTGGAGAGGTGACAAAGGCTGACTGACCGCTGCTACTCACATTCCTTGTGTCAGTCAG 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
QY 443 CACTGTGACTTCTTTTATAAGTATAGTAAACCAAGAAGATCCACACTGCTGTCCTTT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAAGCTATCAAGTCTCGAAGTACCCCAAGTCCCAACCAAGAGGATTTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTTCTTTTGA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGGATTTACATGTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
QY 683 AGAATAATCTCTACTGCTGCTGTTTCATATATAGTGGTCTCTCAGGTGCAGTTCGTAATA 742
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsn 663
QY 743 CGGTTTAAAGAGTCAAGTATTTCCGTTTGAATACCTTACCTAGCCTCTCTAGCTTATGTGTT 802
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
QY 803 GTAGTGATAGACAAACAGGGGATCTCTGTCAACGAGGGCTTAAATTTGAAGCGCTTTAAA 862
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
QY 863 TATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCT 922
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
QY 923 CGATATGATTTCAVTCAGTTCAGTTCGTCGTCATCCAGGTGCAGGTGCTTATGAGGATAC 982
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr 743
QY 983 CTCTCCCTGATGGCATTAAATGCAGAGTCAAGATATCTTCAGGTTCTTCTGCTGGGCC 1042
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
QY 1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081
Db 764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776

```

# RESULT 6

```

US-10-072-012-622
; Sequence 622 Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Beha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie

```

```
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Qusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Molenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 622
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-622

Alignment Scores:
Pred. No.: 9,51e-182 Length: 882
Score: 1836.50 Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8% Indels: 173
DB: 4 Gaps: 1

US-10-825-632-8 (1-1083) x US-10-072-012-622 (1-882)
Qy 2 GAAGAAGATCCAGATCAGCTGAGTCTGCTACCTTTGTTCTCCAGAAGAATTGTGATAGA 61
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263
Qy 62 TATTCGGCTATTGGTGGTCCAAAAGCTGAAACACTCCACAGTGGTGGTAAATTCCT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
Qy 122 AGAATTCATATGAGAAATATGATCTGAGGTGGAAATATTTTCATGTTTACATCCCT 181
Db 284 ArgIleuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303
Qy 182 ATGTGGAACACAGGAGGCGAGATTCAATCCGTTATCCATAAACAGGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
Qy 242 AAGTCACTTTAAGATGTCAGAAATATGATGATGCTGAAGGAGGATCAT- - - - - 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal 343
```

```
Qy 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
Qy 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
Qy 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspVal 403
Qy 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
Qy 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
Qy 295 ----- 295
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
Qy 295 ----- 295
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483
Qy 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
Qy 296 ----- GTTGATGAAGTCAGAAAGCTGGTATAT 322
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
Qy 323 TTTGAAGGCACCAAGACTCCCTTTAGAGACTACCTGTACTGATGTCAGTTACTGTAAT 382
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
Qy 383 CCTGGAGAGGTGACAAAGGCTGACTGACCGTGCTACTCACATTTCTGTCAGTCAGTCAG 442
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563
Qy 443 CACTGTGACTCTTTTATAAGTATAGTATAGTAAACCAAGAGATCCACACTGTGTCCCTT 502
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
Qy 503 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
Qy 563 ATTTGTGATTCAGAGGTCCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623
Qy 623 AGTACTACTCGATTTTACATTGATGGGATGCTCTCAAGCCCTCATGATCTACAGCCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Qy 683 AAGAAATATCCTACTGTGCTGTTTCAATATATGTTGGTCTCTCAGGTCAGTGGTGAATAT 742
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
Qy 743 CGGTTTAAAGAGCTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGTT 802
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
Qy 803 GTAGTGATAGACACAGGGGATCCTGTCCACGAGGCTTAAATTTGAAGGCGCCTTTAAA 862
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
```



Db	744	LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763
Qy	1043	CCAGTCACTCTGTGATCTTATCATACAGATACAGACAG	1081
Db	764	ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr	776
RESULT 8			
US-10-825-632-1			
; Sequence 1, Application US/10825632			
; Publication No. US20040191826A1			
; GENERAL INFORMATION:			
; APPLICANT: ABBOTT, Catherine Anne			
; APPLICANT: GORRELL, Mark Douglas			
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES			
; FILE REFERENCE: FCSB-100-Div. 1			
; CURRENT APPLICATION NUMBER: US/10/825,632			
; CURRENT FILING DATE: 2004-04-15			
; PRIOR APPLICATION NUMBER: US 10/070,464			
; PRIOR FILING DATE: 2002-07-18			
; PRIOR APPLICATION NUMBER: PCT/AU00/01085			
; PRIOR FILING DATE: 2000-09-11			
; PRIOR APPLICATION NUMBER: AU P05709			
; PRIOR FILING DATE: 2000-02-18			
; PRIOR APPLICATION NUMBER: AU P02762			
; PRIOR FILING DATE: 1999-09-10			
; NUMBER OF SEQ ID NOS: 31			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 1			
; LENGTH: 882			
; TYPE: PRT			
; ORGANISM: Homo Sapiens			
US-10-825-632-1			
Alignment Scores:			
Pred. No.:	9,51e-182	Length:	882
Score:	1836.50	Matches:	360
Percent Similarity:	67.5%	Conservative:	0
Best Local Similarity:	67.5%	Mismatches:	0
Query Match:	94.8%	Indels:	173
DB:	4	Gaps:	1
US-10-825-632-8 (1-1083) x US-10-825-632-1 (1-882)			
Qy	2	GAAGAAGATGCCAGATCAGCTGAGTGGCTGCTACCTTGTCTTCCCAAGAAATTTGATAGA	61
Db	244	GlulGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg	263
Qy	62	TATTCGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTCCT	121
Db	264	TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu	283
Qy	122	AGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCT	181
Db	284	ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro	303
Qy	182	ATGTGGAAACAGGAGGCGAGATTCCTCGTTATCTTAAACAGGTACAGCAAAATCCT	241
Db	304	MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro	323
Qy	242	AAAGTCATCTTTAAGATGTCAGAAATATGATTGATGCTGAAGGAGGATCATA-----	295
Db	324	LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspVal	343
Qy	295	-----	295
Db	344	IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla	363
Qy	295	-----	295
Db	364	ArgAlaGlyTrpThrProGluGlyLysTyrAlaThrSerIleLeuLeuAspArgSerGln	383
Qy	295	-----	295
Db	384	ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal	403
Qy	295	-----	295
Db	404	MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr	423
Qy	295	-----	295
Db	424	GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer	443
Qy	295	-----	295
Db	444	HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu	463
Qy	295	-----	295
Db	464	TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro	483
Qy	295	-----	295
Db	484	AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp	503
Qy	296	-----GTTGATGAAGTCAGAGGCTGTGATATAT	322
Db	504	GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr	523
Qy	323	TTTGAAGGCACAAAGACTCCCTTTTAGAGCATCATCTGTAGTGTAGTTACGTAAT	382
Db	524	PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn	543
Qy	383	CCTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGT	442
Db	544	ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln	563
Qy	443	CACTGTGACTCTTTATAAGTATAGTATAACAGAGAATCCACACTGTGTGTCCCTT	502
Db	564	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu	583
Qy	503	TACAAAGCTATCAAGTCTCGAAGATGACCACTTGGCAAAACAAGAAATTTGGGCGACC	562
Db	584	TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr	603
Qy	563	ATTTTGGATTGACAGAGTCTCTTCTGACTATATCTCTCCAGAAATTTTCTTTTGAA	622
Db	604	IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu	623
Qy	623	AGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAGCCTCATGATCTACAGCTTGA	682
Db	624	SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly	643
Qy	683	AAGAAATATCTTACTGTGTGTTTATATATATGTTGGTCTCTCAGGTGAGTGGTGAATAAT	742
Db	644	LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn	663
Qy	743	CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGTATGTGGTT	802
Db	664	ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal	683
Qy	803	GTAGTGTATGATACAAACAGGAGGATCTCTGTCACCGAGGCTTAAATTTGAGGCGCTTAAA	862
Db	684	ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys	703
Qy	863	TATAAATGGTCAAAATAGAAATTCAGATCAGGTGGAGGACTCCATATATCTAGTTCT	922
Db	704	TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer	723
Qy	923	CGATATCATTTTCATTTGACTTTAGATCGTGTGGGATCCACGGCTGTCTCTATGGAGTAC	982
Db	724	ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyr	743
Qy	983	CTCTCCCTGATGGCATTAATTCAGAGGTGAGATATCTTCAGGTTGCTTATTCCTGGGCCC	1042
Db	744	LeuSerMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla	763

QY 1043 CCAGTCACTGTGGTATCTTCTATGATACAGGATACACG 1081  
Db ||||||||||||||||||||||||||||||||||||||||  
764 ProValThrLeuTrpIlePheTyrAspThrGlyTyrThr 776  
RESULT 9  
US-10-982-512-1  
; Sequence 1, Application US/10982512  
; Publication No, US2005059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Qi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV  
; FILE REFERENCE: 70669  
; CURRENT APPLICATION NUMBER: US/10/982,512  
; CURRENT FILING DATE: 2004-11-05  
; PRIOR APPLICATION NUMBER: US/09/976,674  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 60/240,117  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 882  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-982-512-1

Alignment Scores:  
Pred. No.: 9,51e-182 Length: 882  
Score: 1836.50 Matches: 360  
Percent Similarity: 67.5% Conservative: 0  
Best Local Similarity: 67.5% Mismatches: 0  
Query Match: 94.8% Indels: 173  
DB: 5 Gaps: 1

US-10-825-632-8 (1-1083) x US-10-982-512-1 (1-882)

QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGA 61  
Db 244 GluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPheAspArg 263  
QY 62 TATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTT 121  
Db 264 TyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283  
QY 122 AGAATTCATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCT 181  
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerPro 303  
QY 182 ATGTTGAAACAAGAGGCGCAGATTCATTCCGTTATCTTAAACAGGTACAGCAATCCT 241  
Db 304 MetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323  
QY 242 AAAAGTCATTTTAAAGATGTCAAAATAATGATGCTGCTGAAGGAGGATCATA----- 295  
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluArgIleIleAspVal 343  
QY 295 ----- 295  
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363  
QY 295 ----- 295  
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuAspArgSerGln 383  
QY 295 ----- 295  
Db 384 ThrArgLeuGlnIleValIleSerProGluLeuPheIleProValGluAspVal 403  
QY 295 ----- 295

Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423  
QY 295 ----- 295  
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443  
QY 295 ----- 295  
Db 444 HisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463  
QY 295 ----- 295  
Db 464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuPro 483  
QY 295 ----- 295  
Db 484 AlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSerGlyGluTrp 503  
QY 296 -----GTTGATGAAGTCAGAAGCGCTGTATAT 322  
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523  
QY 323 TTGAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTACGTAAAT 382  
Db 524 PheGluGlyThrLysAspSerProLeuIleHisLeuTyrValValSerTyrValAsn 543  
QY 383 CCTGGAGAGGTGACAAAGCTGACTGACCTGGCTACTCACATTTCTTGCTGCATCAGTAC 442  
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGln 563  
QY 443 CACTGTGATCTTTTATTAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGCTTTTGA 622  
Db 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
QY 503 TACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAGAAATTTTGGGCGCAC 562  
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603  
QY 563 ATTTTGGATTCAGCAGGTCTCTTCTGCTGATCTCTCTGCTGCTCTCCAGAAATTTCTTTTGA 622  
Db 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
QY 623 AGTACTACTGATTTACATTGATGGGATCTCTCAAGCCTCATGATCTACAGCTCGA 682  
Db 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
QY 683 AAGAAATATCTACTGTGCTGTTTCATATAAGTGGTCTCTCAGTGCAGTGTGTAATAAT 742  
Db 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663  
QY 743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTATGTGGTT 802  
Db 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683  
QY 803 GTAGTGTATAGACAACAGGGGATCTCTGACCGAGGCTTAAATTTGAAGGCGCTTTTAAA 862  
Db 684 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703  
QY 863 TATAAAATGGTCAATAGAAATTCACGATCAGGTGGAGGACTCCAAATATCTAGCTTCT 922  
Db 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723  
QY 923 CGATATGATTTTCATTGACTAGCTGTGCGGATCCACGCTCGTCTCTATGAGGATAC 982  
Db 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTyr 743  
QY 983 CTCTCCCTGATGATTAATCAGAGGTACAGATATCTTCAGGGTGTGATTTGCTGGGGCC 1042  
Db 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763  
QY 1043 CCAGTCACTGTGGATCTTCTATGATACAGGATACACG 1081  
|||||



```
Db 605 ILeuAspSerAlaGlyProLeuProAspTyrThrProGluLeuPheSerPheGlu 624
Qy 623 AGTACTACTGGATTACATTGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 682
Db 625 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 644
Qy 683 AAGAAATATCTACTGCTGCTGCTCATATATGCTGCTCAGTGCAGTGGCAATAT 742
Db 645 LysLysTyrProThrValLeuPheLeuTyrGlyGlyProGlnValGlnLeuValAsnAsn 664
Qy 743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGATATACCTAGCCTCTCTAGTGTATGGTT 802
Db 665 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 684
Qy 803 GTAGTCATAGACAAACAGGGATCTCTCCACGAGGGCTTAAATTTGAAGCGCCTTTAAA 862
Db 685 ValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 704
Qy 863 TATAAATGGTCAATAGAAATTCAGCATCAGTGGGAAGGACTCCAATATCTAGCTTCT 922
Db 705 TyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 724
Qy 923 CGATATGATTCATTGACTTATAGATCGTGTGGGCATCCAGCGCTGGTCCCTATGGAGGATAC 982
Db 725 GlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 744
Qy 983 CTCCTCCTGATGGCAATTAATCCAGAGTCAATATCTTCAGGGTTCCTATTCCTGGGCC 1042
Db 745 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 764
Qy 1043 CAGTCATCTGTGGATCTTCTATGATACAGGATACAGC 1081
Db 765 ProValThrLeuThrPheTyrAspThrGlyTyrThr 777

RESULT 11
US-10-275-505-2
; Sequence, 2, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DELEGANE, Angelo M.; LAL, Preeti G.
; APPLICANT: HAPALIA, April J.A.; PATTERSON, Chandra
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.
; APPLICANT: NGUYEN, Daniel B.; GANDHI, Ameena R.
; APPLICANT: YANG, Junming; HERNANDEZ, Roberto
; APPLICANT: POLICKY, Jennifer L.; LU, Dyung Aina M.
; APPLICANT: REDDY, Roopa M.; YUE, Henry
; APPLICANT: TANG, Y. Tom
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0085 USN
; CURRENT APPLICATION NUMBER: US/10/275,505
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US01/14651
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/209,402
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/207,477
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/205,803
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/203,566
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/202,082
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 580
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 376067CD1
US-10-275-505-2

Alignment Scores:
Pred. No.: 3,02e-146 Length: 580
Score: 1495.00 Matches: 306
Percent Similarity: 57.8% Conservative: 0
Best Local Similarity: 57.8% Mismatches: 1
Query Match: 77.1% Indels: 223
DB: 4 Gaps: 2

US-10-825-632-8 (1-1083) x US-10-275-505-2 (1-580)
Qy 9 ATCCGAGATCAGTGGAGTCGCTACCTTGTCTCTCCAGAGAATTTGATAGATATTCTG 68
Db 1 MetProAspGlnLeuGlnSerLeuProLeuPheSerLysLysAsnLeuIleAspIleLeu 20
Qy 69 GCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCACTGGTGGTAAATTTCTTGAATTC 128
Db 21 AlaIleGlyValGlnLeuLysLeuLysGlnLeuProValValLysPheLeuGluPhe 40
Qy 129 TATATGAGAAAATGATGAATCTGAGTGGAAATTTATTCATGTTCATCCCTATGTGG 188
Db 41 TyrMetLysLysMetMetAsnLeuArgTrpLysLeuPheMetLeuHisProLeuCysTrp 60
Qy 189 AAACAGGAGG- GCAGATTTCATTCCTTATCCTTAAACAGGTACAGCAATCTCTAAGTC 247
Db 61 LysGlnGlyArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnProLysVal 80
Qy 248 ACTTTAAAGATGTCAGAAATAATGATTGATCGCTGAAGGAGGATCAT- - - - - 295
Db 81 ThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleIleAspValIleAsp 100
Qy 295 - - - - - 295
Db 101 LysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAlaArgAla 120
Qy 295 - - - - - 295
Db 121 GlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArg 140
Qy 295 - - - - - 295
Db 141 LeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspValMetGlu 160
Qy 295 - - - - - 295
Db 161 ArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyrGluGlu 180
Qy 295 - - - - - 295
Db 181 ThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSerHisGlu 200
Qy 295 - - - - - 295
Db 201 GluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeuTyrLys 220
Qy 295 - - - - - 295
Db 221 IleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuProAlaPro 240
Qy 295 - - - - - 295
Db 241 ThrValThrTrpMetIleThrPheMetArgSerLeuGlyThrProSerCysMetCysVal 260
Qy 296 - - - - - GTTGATGAAGTCAAGAGCGCTGATATTTTGAAGGCACC 334
Db 261 ThrHisIleValGluIleGlnValAspGluValArgArgLeuValTyrPheGluGlyThr 280
Qy 335 AAAGACTCCCTTTAGAGCATCATCCTGACGTACGTACGTTACGTTAAATCCTCGAGAGGTG 394
```



Db 281 LysAspSerProLeuGluHisLeuTyrValValSerTyrValAsnProGlyGluVal 300  
QY 395 ACAGGCTGACTGACGGTGGTACTACATCTTGGCTGCATCAGTCACTGACTTC 454  
Db 301 ThrArgLeuThrAspArgGlyTyrSerHisSerCysHisSerGlnHisCysAspPhe 320  
QY 455 TTTATAAGTAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTTACAAAGCTATCA 514  
Db 321 PheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSer 340  
QY 515 AGTCTGAAGATGACCAACTTGCAAAACAAGAAATTTGGCCACCATTTGGATTCA 574  
Db 341 SerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSer 360  
QY 575 GCAGGTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGA 634  
Db 361 AlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSerThrThrGly 380  
QY 635 TTTACATTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTCT 694  
Db 381 PheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrPro 400  
QY 695 ACTGTGCTGTTCATATATGCTGCTCTCAGGTGCAGTTGGTGAATAATCGTTTAAAGGA 754  
Db 401 ThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGly 420  
QY 755 GTCAAGTATTTCCGTTGAATACCTAGCCTCTAGGTGTAAGTGTAGTGTAGTATGAC 814  
Db 421 ValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValValValIleAsp 440  
QY 815 AACAGGGATCTGTCCACGGCGCTTAAATTTGAAGCGCTTTAAATATAAATGGGT 874  
Db 441 AsnArgLysSerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMet 459  
QY 875 CAATAGAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTC 934  
Db 459 ----- 459  
QY 935 ATTGACTTAGATCGTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATG 994  
Db 459 ----- 459  
QY 995 GCATTAAATGACAGGTGAGATATCTTACGGTTGCTATGCTGGGGCCCGACGTCTG 1054  
Db 460 -----ValAlaIleAlaGlyAlaProValThrLeu 469  
QY 1055 TGAATCTTCTATGATACAGGATACAG 1081  
Db 470 TrpIlePheTyrAspThrGlyTyrThr 478

RESULT 12

US-11-140-224-2  
; Sequence 2, Application US/11140224  
; Publication No. US2005027280A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: DELEGEANE, Angelo M.; LAL, Preeti G.  
; APPLICANT: HAFALIA, April J.A.; PATTERSON, Chandra  
; APPLICANT: WALIA, Narinder K.; KEARNEY, Liam  
; APPLICANT: TRIBOULEY, Catherine M.; KHAN, Farrah A.  
; APPLICANT: YAO, Monique G.; BAUGHN, Mariah R.  
; APPLICANT: AZIMZAI, Yalda; ELLIOTT, Vicki S.  
; APPLICANT: NGUYEN, Daniel B.; GANDHI, Ameena R.  
; APPLICANT: YANG, Junming; HERNANDEZ, Roberto  
; APPLICANT: POLICKY, Jennifer L.; LU, Dying Aina M.  
; APPLICANT: REDDY, Roopa M.; YUE, Henry  
; APPLICANT: TANG, Y. Tom  
; TITLE OF INVENTION: PROTEASES  
; FILE REFERENCE: FI-0085 USN  
; CURRENT APPLICATION NUMBER: US/11/140,224  
; CURRENT FILING DATE: 2005-05-31  
; PRIOR APPLICATION NUMBER: US/10/275,505

; PRIOR FILING DATE: 2002-11-04  
; PRIOR APPLICATION NUMBER: PCT/US01/14651  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/209,402  
; PRIOR FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: 60/207,477  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/205,803  
; PRIOR FILING DATE: 2000-05-17  
; PRIOR APPLICATION NUMBER: 60/203,566  
; PRIOR FILING DATE: 2000-05-11  
; PRIOR APPLICATION NUMBER: 60/202,082  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PERL Program  
; SEQ ID NO 2  
; LENGTH: 580  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 376067CD1  
US-11-140-224-2  
Alignment Scores:  
Pred. No.: 3,02e-146 Length: 580  
Score: 1495.00 Matches: 306  
Percent Similarity: 57.8% Conservatives: 0  
Best Local Similarity: 57.8% Mismatches: 1  
Query Match: 77.1% Indels: 223  
DB: 6 Gaps: 2  
US-10-825-632-8 (1-1083) x US-11-140-224-2 (1-580)  
QY 9 ATGCCAGATCAGCTGGAGTCGCTACCTTTTGTCTCCAAAGAAATTTGTAGATATCTG 68  
Db 1 MetProAspGlnLeuGluSerLeuProLeuPheSerLysLysAsnLeuIleAspIleLeu 20  
QY 69 GCTATTGCTGTCTCCAAAGAGCTGAAACAACTCCACAGTGGTGGTAAAAATTTCTTAGAATTC 128  
Db 21 AlaIleGlyValGlnLysLeuLysGlnLeuProValValLysPheLeuGluPhe 40  
QY 129 TATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTGG 188  
Db 41 TyrMetLysLysMetMetAsnLeuArgTrpLysLeuPheMetLeuHisProLeuCysTrp 60  
QY 189 AAACAAGGAGG-GCAGATTTCATTCCTGTTATCTTAAACAGGTACAGCAAAATCCTAAAGTC 247  
Db 61 LysGlnGlyArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnProLysVal 80  
QY 248 ACTTTTAAGATGTCAGAAATAATGATGATGCTGGAAGGAGGATCATATA----- 295  
Db 81 ThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleAspValIleAsp 100  
QY 295 ----- 295  
Db 101 LysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAlaArgAla 120  
QY 295 ----- 295  
Db 121 GlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArg 140  
QY 295 ----- 295  
Db 141 LeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspValMetGlu 160  
QY 295 ----- 295  
Db 161 ArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyrGluGlu 180  
QY 295 ----- 295  
Db 181 ThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSerHisGlu 200

```
QY 295 ----- 295
Db 201 GluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheA-rgHisLeuTyrLys 220
QY 295 ----- 295
Db 221 IleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuProAlaPro 240
QY 295 ----- 295
Db 241 ThrValThrTrpMetIleThrPheMetArgSerLeuGlyThrProSerCysMetCysVal 260
QY 296 ----- 296
Db 261 ThrHisIleValGluIleGlnValAspGluValArgArgLeuValTyrPheGluGlyThr 280
QY 335 AAAGACTCCCTTTAGACATCACCTGTACGTAGTACGTAGTACGTAACTCTGAGAGGNG 394
Db 281 LysAspSerProLeuGluHisLeuTyrValValSerTyrValAsnProGlyGluVal 300
QY 395 ACAAGCTGATGACCGTGGCTACTCATCTTGTCTGTCATCATGTCAGCTGTCACCTTC 454
Db 301 ThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGlnHisCysAspPhe 320
QY 455 TTTATAAGTATAGTAAACAGAGAATCCACACTGTGTCTGCTTACAGCTATCA 514
Db 321 PheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSer 340
QY 515 AGTCTGAAGATGACCCAACTTGCNAACAAAGGAATTTGGGCCACCATTTGGATTCA 574
Db 341 SerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSer 360
QY 575 GCAGGTCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGGA 634
Db 361 AlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSerThrGly 380
QY 635 TTTACATGTATGGGATGCTACACAGCCTCATGCTACAGCCTGGAAGAATAATPCT 694
Db 381 PheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrPro 400
QY 695 ACTGTGCTTTCATATATCGTCTCAGTCTCAGTCTGCTGCTGCTGCTGCTGCTGCT 754
Db 401 ThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGly 420
QY 755 GTCAACTATTTCGCTTGATACCTAGCCTCTCTAGGTTATGCTGTTAGTGTAGTAGAC 814
Db 421 ValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValValValIleAsp 440
QY 815 AACAGGGATCTGTACAGGAGGCTTAAATTTCAAGCGCCTTTAAATATATAAATGGGT 874
Db 441 AsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMet 459
QY 875 CAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTC 934
Db 459 ----- 459
QY 935 ATTGACTTAGATCGTGTGGGATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATG 994
Db 459 ----- 459
QY 995 GCATTATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTG 1054
Db 460 ----- 460
QY 1055 TGGATCTTCTATGATACAGGATACAG 1081
Db 470 TrpIlePheTyrAspThrGlyTyrThr 478
```

## RESULT 13

US-09-976-674-7

; Sequence 7, Application US/09976674

; Patent No. US20020115843A1

```
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-7
```

## Alignment Scores:

Pred. No.:	3,246-118	Length:	690
Score:	1226.00	Matches:	259
Percent Similarity:	48.5%	Conservative:	0
Best Local Similarity:	48.5%	Mismatches:	0
Query Match:	63.3%	Indels:	275
DB:	3	Gaps:	2

US-10-825-632-8 (1-1083) x US-09-976-674-7 (1-690)

```
QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAATTTGATAGA 61
Db 244 GluGluAspAlaAspSerArgLysValAlaThrPheValLeuGlnGluPheAspArg 263
QY 62 TATTCTCGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAAAATCTT 121
Db 264 TyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGlyLysIleLeu 283
QY 122 AGAATTCATATGAAGAAATGATGATCTGAGGTGGAAATTTATTCATGTTACATCCCT 181
Db 284 ArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisValThrSerPro 303
QY 182 ATGTTGMAACAAGAGGCGCAGATTCATTCCTGTTATCTCTAAACAGGTACAGCAATCCT 241
Db 304 MetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnPro 323
QY 242 AAAGTCATTTTAAGATGTCAAGAAATAATGATTGATGCTGAAGGAAGGATCATA---- 295
Db 324 LysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyA-gIleIleAspVal 343
QY 295 ----- 295
Db 344 IleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAla 363
QY 295 ----- 295
Db 364 ArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGln 383
QY 295 ----- 295
Db 384 ThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspAspVal 403
QY 295 ----- 295
Db 404 MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyr 423
QY 295 ----- 295
Db 424 GluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSer 443
QY 295 ----- 295
Db 444 HisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeu 463
```

us-10-825-632-8.n2p.rapbm

Mon May 8 13:41:59 2006

TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

FILE REFERENCE: 70669

CURRENT APPLICATION NUMBER: US/10/982,512

CURRENT FILING DATE: 2004-11-05

PRIOR APPLICATION NUMBER: US/09/976,674

PRIOR FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: US 60/240,117

PRIOR FILING DATE: 2000-10-12

NUMBER OF SEQ ID NOS: 61

SOFTWARE: Patent in version 3.1

SEQ ID NO 7

LENGTH: 690

TYPE: PRT

ORGANISM: Homo sapiens

US-10-982-512-7

Alignment Scores: 3.24e-118 Length: 690  
Pred. No.: 1226.00 Matches: 259  
Score: 1226.00 Conservative: 0  
Percent Similarity: 48.5% Mismatches: 0  
Best Local Similarity: 48.5% Indels: 275  
Query Match: 63.3% Gaps: 2  
DB: 5

US-10-825-632-8 (1-1083) x US-10-982-512-7 (1-690)

295 ----- 295  
464 TyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyLeuPro 483  
295 ----- 295  
484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTyr 503  
296 ----- 296  
504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyr 523  
323 TTTGAAGGCACAAAGACTCCCTTTAGACCATCCTGTACGTAGTACGTAAAT 382  
524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543  
383 CTGGAGAGGTGACAAAGCTGACGCTGGCTACTCACATTTCTGTCATGTCAG 442  
544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563  
443 CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAATCCACACTGTGTCCCTT 502  
564 HisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583  
503 TACAAGCTATCAAGTCTCGAAGATGACCCAACTTGCACAAACAAAGGAATTTGGGCCACC 562  
584 TyrLysLeuSerSerProGluAspSerProThrCysLysThrLysGluPheTyrPalaThr 603  
563 ATTTTGGATTACAGAGTCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTGAA 622  
604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePheSerPheGlu 623  
623 AGTACTACTGGATTACATTTATGGAGTCTTCAATATATGGTGGTCTCAGTGTGGA 682  
624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643  
683 AAGAAATATCTACTGTGCTGTTTCATATATATGGTGGTCTCAGTGTGGA 742  
644 LysLysTyrProThrValLeuPheIleTyrGlyGly----- 655  
743 CGGTTTAAAGGAGTCAAGTATTTCCCGTTGAATACCTTAGCTCTCTAGTTATGTGTT 802  
655 ----- 655  
803 GTAGTGATAGACACAGGGATCCTGTACCGAGGCTTAAATTTGAAGGCCCTTTAAA 862  
655 ----- 655  
863 TATAAATGGGTCAAAATAGAAATTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCT 922  
655 ----- 655  
923 CGATATGATTTCAITGACTTAGATCGTGTGGGCATCCACGCTGCTCTATGGAGGATAC 982  
655 ----- 655  
983 CTCTCCCTGATGGCAATTAATGCAGAGGTACAGATATCTTACGGTTGCTATTCTGGGGCC 1042  
656 ----- 656  
1043 CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGG 1082  
662 OGInSerCysGlySerSerMetIleGlnAspThrArg 675

RESULT 14

US-10-982-512-7  
; Sequence 7, Application US/10982512  
; Publication No. US20050059081A1  
; GENERAL INFORMATION:  
; APPLICANT: Oi, Steve  
; APPLICANT: Akinsanya, Karen  
; APPLICANT: Riviere, Pierre  
; APPLICANT: Junien, Jean-Louis

```
QY 295 ----- 295
Db 484 AlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrp 503
QY 296 ----- 296
Db 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluVala-GaGleuValTrp 523
QY 323 TTTGAAGGCCCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGTCAGTATAT 322
Db 524 PheGluGlyThrLysAspSerProLeuGluHisLysLeuTrpValValSerTrpValAsn 543
QY 383 CTTGGAGAGTCAAGGCTGACTGACCGTGGCTACTCACATCTTGTGTCATCAGTCAG 382
Db 544 ProGlyGluValThrArgLeuThrAspArgGlyTrpSerHisSerCysValSerGln 563
QY 443 CACTGTGACTCTTTTAAAGTATAGTATAGTAAAGATGCCCACTTGCAAAACAAAGAA 562
Db 564 HisCysAspPhePheIleSerLysTrpSerAsnGlnLysAsnProHisCysValSerLeu 583
QY 503 TACAGCTATCAACTCTGAAGATGACCCCACTTGCAAAACAAAGAAATTTTGGGCCACC 562
Db 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
QY 563 ATTTTGGATTACAGAGTCTCTTCTGACTACTCTTCCAGAAATTTTCTTTTGA 622
Db 604 IleLeuAspSerAlaGlyProLeuProAspTrpThrProGluIlePheSerPheGlu 623
QY 623 AGTACTACTGATTATCATGGATGGATGCTCTCAAGCCTCATGATCTACAGCCTGGA 682
Db 624 SerThrThrGlyPheThrLeuTrpGlyMetLeuTrpLysProHisAspLeuGlnProGly 643
QY 683 AAGAAATATCTACTGTGCTGTATATATGTTGCTCTCAGGTGCAGTGGTGAATAAT 742
Db 644 LysLysTrpProThrValLeuPheIleTrpGlyGly----- 655
QY 743 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTTAGGTTATG 802
Db 655 ----- 655
QY 803 GTAGTGATAGACAAAGGGGATCCTGTCAACGGGGCTTAATTTGAAGCGCCTTTAAA 862
Db 655 ----- 655
QY 863 TATAAAATGGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTCT 922
Db 655 ----- 655
QY 923 CGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTATGAGGATAC 982
Db 655 ----- 655
QY 983 CTCTCCCTGATGGCATTAAATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGCC 1042
Db 656 ----- 656
QY 1043 CCAGTACTCTGTGGATCTTCTATATGATACAGGATACACGG 1082
Db 662 OGInSerLeuGlySerSerMetIleGlnAspThrArg 675

RESULT 15
US-10-072-012-620
; Sequence 620, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tcheurev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Bsha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Restelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier, J. Raymond J.
; APPLICANT: Guev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Molenc, Adolphe A.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-620
Alignment Scores:
Pred. No.: 1,21e-117 Length: 689
Score: 1220.50 Matches: 689
Percent Similarity: 54.3% Conservative: 54
Best Local Similarity: 44.2% Mismatches: 69
Query Match: 63.0% Indels: 175
DB: 4 Gaps: 2
US-10-825-632-8 (1-1083) x US-10-072-012-620 (1-689)
QY 5 GAAGATCCAGATCAGTCGAGTCGCTACCTTTGTTCTCCAGAAAGATTTGTATAGATAT 64
Db 60 AspAspProLysSerAlaGlyValAlaThrPheValIleGlnGluPheAspArgPhe 79
QY 65 TCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACACTCCACCTGGTGGT---AAAATCTT 121
Db 80 ThrGlyTrpTrpTrpCysProThrAlaSerTrpGluGlySerGluGlyLeuThrLeu 99
QY 122 AGAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATGTTTACATCCCT 181
Db 100 ArgIleLeuTrpGluGluValAspGluSerGluValGluValIleHisValProSerPro 119
QY 182 ATGTTGGAACAAAGGAGGCGAGATTTCATTCGTTTCTTAAACAGGTACAGCAATCT 241
; APPLICANT: Gangolli, Bsha
```

```
Db 120 AlaLeuGluGluArgLysThrAspSerTyrArgTyrProArgThrGlySerLysAsnPro 139
QY 242 AAGTCACCTTTAAGATGTCAGAAATATGATGCTGTAAGCAGATCATA----- 295
Db 140 LysIleAlaLeuLysLeuAlaGluPheGlnThrAspSerGlnGlyLysIleValSerThr 159
QY 295 ----- 295
Db 160 GlnGluLysGluLeuValGlnProPheSerSerLeuPheProLysValGluTyrIleAla 179
QY 295 ----- 295
Db 180 ArgAlaGlyTyrThrArgAspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGln 199
QY 295 ----- 295
Db 200 GlnTrpLeuGlnLeuValLeuLeuProAlaLeuPheIleProSerThrGluAsnGlu 219
QY 295 ----- 295
Db 220 GluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnProTyrValValTyr 239
QY 295 ----- 295
Db 240 GluGluValThrAsnValTrpIleAsnValHisAspIlePheTyrProPheProGlnSer 259
QY 295 ----- 295
Db 260 GluGlyGluAspGluLeuCysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHis 279
QY 295 ----- 295
Db 280 LeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSer 299
QY 295 ----- 295
Db 300 ProGlyGluAspGluPheLysCysProIleLysGluGluIleAlaLeuThrSerGlyGlu 319
QY 296 -----GTTGATGAAGTCAGAGCGTGGA 319
Db 320 TrpGluValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuVal 339
QY 320 TATTTGAAGCACCAGACATCCCTTTAGAGCATCACCTGTAGTAGTACGATTACGTA 379
Db 340 TyrPheGlnGlyThrLysAspThrProLeuGluHisHisLeuTyrValValSerTyrGlu 359
QY 380 AATCCTGGAGAGGTGCACAGGCTGACTGACCGGTGCTACTCACATTCTTCTGTCATCAGT 439
Db 360 AlaAlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSer 379
QY 440 CACACTGTGACTTCTTTAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGTC 499
Db 380 GlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrProProCysValHis 399
QY 500 CTTTACAAGCTATCAAGTCTCAAGATGACCCACTTGCAAAACAAAGGAATTTGGGCC 559
Db 400 ValTyrLysLeuSerGlyProAspAspProLeuHisLysGlnProArgPheTrpAla 419
QY 560 ACCATTTTGGATTTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAATTTCTCTTTT 619
Db 420 SerMetMetGluAlaAlaSerCysProProAspTyrValProProGluIlePheHisPhe 439
QY 620 GAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCT 679
Db 440 HisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnPro 459
QY 680 GGAAGAAATATCCTACTGCTGCTTCATATATATGTTGCTCCTCAGGTGCAGTTGGTGAAT 739
Db 460 GlyLysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn 479
QY 740 AATCGGTTAAAGGAGTCAAGTATTTCCGCTTGATACCTAGCCTCTAGGTATGTG 799
Db 480 AsnSerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAla 499
```

```
QY 800 GTTCTAGTGTAGACACAGGAGTCTCTGTACCGAGGCTTAAATTTGAAGCGCCTTT 859
Db 500 ValValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeu 519
QY 860 AAATATAAATGGGTCAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCT 919
Db 520 LysAsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAla 539
QY 920 TCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGA 979
Db 540 GluLysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGly 559
QY 980 TACCTCTCCCTGATGGCATTAAATGCAGAGTCAAGATATCTTCAGGTTCTTATGCTGGG 1039
Db 560 PheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGly 579
QY 1040 GCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACG 1081
Db 580 AlaProValThrValTrpMetAlaTyrAspThrGlyTyrThr 593
```

Search completed: May 2, 2006, 03:08:58  
Job time : 179.086 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: May 2, 2006, 02:50:19 ; Search time 3.62006 Seconds  
(without alignments)  
2721.092 Million cell updates/sec

Title: US-10-825-632-8  
Perfect score: 1938  
Sequence: 1 ggaagaagatgccagatcag.....tatgatacagatacacgga 1083

Scoring table: BIOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 464238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_n2p.model -DEV=xlp  
-Q=/abes/ABSWEB\_spool/US10825632/runat\_01052006\_110000\_3533/app\_query.fasta 1  
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abs802p  
-USER=US10825632 @CGN\_1\_1\_59 @runat\_01052006\_110000\_3533 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:

- 1: /SIDSS/ptodata/1/pubpaa/US08 NEW PUB pep.:
- 2: /SIDSS/ptodata/1/pubpaa/US06 NEW PUB pep.:
- 3: /SIDSS/ptodata/1/pubpaa/US07 NEW PUB pep.:
- 4: /SIDSS/ptodata/1/pubpaa/US08 NEW PUB pep.:
- 5: /SIDSS/ptodata/1/pubpaa/PCT NEW PUB pep.:
- 6: /SIDSS/ptodata/1/pubpaa/US05 NEW PUB pep.:
- 7: /SIDSS/ptodata/1/pubpaa/US09 NEW PUB pep.:
- 8: /SIDSS/ptodata/1/pubpaa/US10 NEW PUB pep.:
- 9: /SIDSS/ptodata/1/pubpaa/US11 NEW PUB pep.:
- 10: /SIDSS/ptodata/1/pubpaa/US11 NEW PUB pep.:
- 11: /SIDSS/ptodata/1/pubpaa/US11 NEW PUB pep.:
- 12: /SIDSS/ptodata/1/pubpaa/US60 NEW PUB pep.:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1836.5	94.8	882	11	US-11-151-601-20
2	984	50.8	497	11	US-11-264-096-1594
3	980	50.6	497	11	US-11-264-096-1593
4	383	19.8	627	11	US-11-079-463-7758
5	372	19.2	745	11	US-11-079-463-6408

6	315	16.3	738	11	US-11-208-288-4	Sequence 4, Appli
7	315	16.3	762	11	US-11-116-939-13	Sequence 13, Appli
8	315	16.3	766	9	US-10-522-789-2	Sequence 2, Appli
9	315	16.3	766	9	US-10-501-035-234	Sequence 234, App
10	315	16.3	766	11	US-11-208-288-2	Sequence 2, Appli
11	309	15.9	760	11	US-11-208-288-6	Sequence 6, Appli
12	257	13.3	760	11	US-11-186-284-55	Sequence 55, Appli
13	199	10.3	624	11	US-11-079-463-7504	Sequence 7504, Ap
14	196	10.1	657	11	US-11-179-977-1	Sequence 1, Appli
15	194	10.0	349	11	US-11-264-096-1591	Sequence 1591, Ap
16	141.5	7.3	737	11	US-11-079-463-9281	Sequence 9281, Ap
17	127.5	6.6	555	11	US-11-045-004-1565	Sequence 1565, Ap
18	125	6.4	24	11	US-11-176-951-16	Sequence 16, Appli
19	101.5	5.2	298	11	US-11-179-977-6	Sequence 6, Appli
20	101	5.2	318	11	US-11-179-977-16	Sequence 16, Appli
21	100.5	5.2	360	11	US-11-079-463-5396	Sequence 5396, Ap
22	96	5.0	710	11	US-11-151-601-23	Sequence 23, Appli
23	96	5.0	710	11	US-11-288-493-44	Sequence 44, Appli
24	94.5	4.9	803	9	US-10-204-639-44	Sequence 44, Appli
25	93	4.8	24	11	US-11-176-951-17	Sequence 17, Appli
26	92	4.7	2515	11	US-11-113-424-53	Sequence 53, Appli
27	91.5	4.7	849	9	US-10-909-769-18	Sequence 18, Appli
28	91.5	4.7	1067	11	US/11/062	Sequence 3, Appli
29	91.5	4.7	1092	11	US/11/062	Sequence 6, Appli
30	90	4.6	519	11	US-11-079-463-5993	Sequence 5993, Ap
31	87.5	4.5	475	11	US-11-096-568A-31619	Sequence 31619, A
32	87	4.5	441	9	US-10-995-561-638	Sequence 638, App
33	87	4.5	441	9	US-10-995-561-639	Sequence 639, App
34	87	4.5	441	11	US-11-108-519-14	Sequence 14, Appli
35	87	4.5	441	11	US-11-141-554B-14	Sequence 14, Appli
36	86.5	4.5	300	11	US-11-179-977-4	Sequence 4, Appli
37	86.5	4.5	3063	11	US-11-186-284-26	Sequence 26, Appli
38	86	4.4	925	11	US-11-087-099-8562	Sequence 8562, Ap
39	86	4.4	925	11	US-11-188-298-18967	Sequence 18967, A
40	85.5	4.4	816	9	US-10-216-161A-375	Sequence 375, App
41	85	4.4	317	11	US-11-096-568A-16980	Sequence 16980, A
42	85	4.4	332	11	US-11-096-568A-22602	Sequence 22602, A
43	85	4.4	361	11	US-11-096-568A-16979	Sequence 16979, A
44	85	4.4	364	11	US-11-096-568A-22601	Sequence 22601, A
45	85	4.4	402	11	US-11-096-568A-16978	Sequence 16978, A

ALIGNMENTS

RESULT 1  
US-11-151-601-20  
; Sequence 20, Application US/11/151601  
; Publication No. US20060003413A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Weich, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,  
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF  
; FILE REFERENCE: NPI00-054P1RC10M1D1VIM  
; CURRENT APPLICATION NUMBER: US/11/151,601  
; CURRENT FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: US 10/170,789  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/882,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15



```
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Alignment Scores:
Pred. No.: 1,81e-175      Length: 882
Score: 1836.50           Matches: 360
Percent Similarity: 67.5% Conservative: 0
Best Local Similarity: 67.5% Mismatches: 0
Query Match: 94.8%      Indels: 173
DB: 11                    Gaps: 1

US-10-825-632-8 (1-1083) x US-11-151-601-20 (1-882)

QY      2 GAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAATTGTATAGA 61
Db
244  LAGLUUASPALAARGSERALAGLYVALALATHRPHEVALLEUGLUGLUPHEASPARG 263
QY      62 TATCTGCTGATTTGGTGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTT 121
Db
264  TYRSGRGLYTRTPCTPCTPCTPCTPCTPCTPCTPCTPCTPCTPCTPCTPCTPCTPCTPCT 121
QY      122 AGAATTCATATGAGAAATATGAAATCTGAGGTGGAATATTCATCTACATCCCT 181
Db
284  ARGILLELEUTYRGLUGLUASNAEPGLUSERGLUGLULEILEHISVALTHRSPRO 303
QY      182 ATGTTGGAACAAAGAGGAGGAGATTTCATTCCTGATTCCTAAACAGGTACACCAATCCT 241
Db
304  METLEUGLUThrA-GA-GA-ASP-SER-PHE-ARG-TYR-PRO-LYS-THR-GLY-THR-ALA-ASN-PRO 323
QY      242 AAAGTCACCTTTTAAGATGTGCAAAATAATGATGCTGCTGAGGAGGAGGATCATA 295
Db
324  LYSVALTHRPHGLYMETSERGLUILEMETILEASPALAGLUGLYARGILLEILEASPV 343
QY      295 ----- 295
Db
344  ILEASPLYSGLULEULEGlnPROpHEGLUULELEUPHEGLUGLYVALGLUTYRILEALA 363
QY      295 ----- 295
Db
364  ARGALAGLYTRPThrPROGLUGLYLYSTYRAlATrPserILELeuLeuASPARGserGln 383
QY      295 ----- 295
Db
384  THRARGLEUGlnLEValLEULESerPROGLULeUPHEILEProVALGLUASPVal 403
QY      295 ----- 295
Db
404  METGLUARGGlnARGLEULEGLUSerVALProASPserVALThrPROLeuILEILETYR 423
QY      295 ----- 295
Db
424  GLUGLUThrTHRAsPLeITrPLeAsNILEHisASPLeILEPHEHisVALPHEProGlnSer 443
QY      295 ----- 295
Db
444  HISGLUGLULEULEGLUPHEILEPHEALASERGLUCYSLYSTHRGLYPHEARGHISLEU 463
QY      295 ----- 295
```

```
Db      464  TYRLYSILETHRserILELeuLYSGLUSerLYSTYRLeuARGserSerGLYGLYLeuPro 483
QY      295 ----- 295
Db      484  ALAProSerASPpHELYSCYSProLLeLYSGluGluLEAlaILEThrSerGLYGLUTrp 503
QY      296 ----- 296
Db      504  GLUValLeuGLYARGHISGLYSerASNILEGlnValASPGLUValARGLeuValTYR 523
QY      323  TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGTACGTAAAT 382
Db      524  PHEGLUGLYThRLYSASPserProLeuGluHISLeuTYRValValserTYRValASN 543
QY      383  CCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTCAGTCAG 442
Db      544  PROGLYGLUValThrARGLeuThrASPARGGLYTYRserHISserCYSsILEserGln 563
QY      443  CACTGTGACTTCTTTATAGTAAGTATAGTAACAGAGAGATCCACACTGTGTGTCCTT 502
Db      564  HISCYASPpHEPHEILESerLYSTYRserAENGLINLYSAsnPROHISCYsVALserLeu 583
QY      503  TACAAGCTATCAAGTCTCTGAAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGCCACC 562
Db      584  TYRLeuLeuSerSerPROGLUASPpProThrCYSLYSThrLYSGluPHETrPALaThr 603
QY      563  ATTTTGGATTTCAGAGGTCCTTCTCTGACTATATCTCTCCAGAAAATTTTCTTTTGAA 622
Db      604  ILELeuASPserALAGLYProLeuProASPpTYRThrPROProGLUILEPHEserPHEGLU 623
QY      623  AGTACTACTGGATTTTACATTTGATGGAGTGCTCTACAGGCTCATGATCTACAGCCTGGA 682
Db      624  SerThrThrGLYPHEThrLeuTYRGLYMetLeuTYRLYSProHISAspLeuGlnPROGLY 643
QY      683  AAGAAATATCTCTACTGTGCTCTTCCGCTTGAATACCTACAGCTCTCTAGCTTATGGTT 742
Db      644  LYSLYSTYRProThrValLeuPHEILETYRGLYGLYProGlnValGlnLeuValASN 663
QY      743  CGGTTTAAAGGAGCAAGATTTTCCGCTTGAATACCTACAGCTCTCTAGCTTATGGTT 802
Db      664  ARGpHELYSGLYValLYSTYR-PHEARGLeuASNThrLeuAlaSerLeuGLYTYRValVal 683
QY      803  GTAGTGATAGACACAGGGATCTGTACCGAGGGCTTAAATTTGAAGCGCCCTTTAA 862
Db      684  VALValILEASPAsnARGLYSerCYsHISARGGLYLeuLYSPHEGLUGLYAlaPHELYS 703
QY      863  TATAAAATGGGTCAAATAGAAATTCAGCATCAGGTGGAAGGACTCCCAATATCTAGCTTCT 922
Db      704  TYRLYSMetGLYGLINILEGLUILEASPpGlnValGLUGLYLeuGlnTYRLeuAlaSer 723
QY      923  CGATATGATTTTCATTTAGATCTGTGGGCATCCAGCGCTGCTCTATCGAGGATAC 982
Db      724  ARGTYRASPpHEILEASPLeuASPARGValGLYILEHISGLYTrpSerTYRGLYTYR 743
QY      983  CTCTCCCTGATGGCATTATGACAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGCC 1042
Db      744  LeuSerLeuMetAlaLeuMetGlnARGserAsPLePHEARGValAlaILEAlaGLYAla 763
QY      1043  CCAGTCACTCTGTGGATCTTCTATGATACAGGATACAG 1081
Db      764  ProValThrLeuTrpILEPHETYRASPpThrGLYTYRThr 776
```

## RESULT 2

```
US-11-264-096-1594
; Sequence 1594, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS46D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
```

```

; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1594
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-264-096-1594

Alignment Scores:
Pred. No.: 3,23e-90 Length: 497
Score: 984.00 Matches: 177
Percent Similarity: 79.5% Conservative: 40
Best Local Similarity: 64.8% Mismatches: 56
Query Match: 50.8% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-8 (1-1083) x US-11-264-096-1594 (1-497)
QY 263 GAAATAATGATTGATGCTGAAGAGGATCATAGTTCAGAGGCTGGTATAT 322
Db 129 GluValLeuAlaArgHisGlySerLysIleTrpValAsnGluThrLysLeuValTyr 148
QY 323 TTGGAAGCACCAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAAT 382
Db 149 PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAla 168
QY 383 CCGTGAGAGGTGACAGGCTGACTGACCGTGGCTACTCATCTTGTGTCATGTCAG 442
Db 169 AlaGlyGluIleValArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGln 188
QY 443 CACTGTGACTCTTTTATAAGTAAGTATAGTAACACAGAGAAATCCACACTGTGCTCCTT 502
Db 189 AsnPheAspMetPheValSerHisTyrSerValSerThrProCysValHisVal 208
QY 503 TACAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCACAAACAAAGGAATTTTGGGCCACC 562
Db 209 TyrLysLeuSerGlyProAspAspProLeuHisGlyGlnProArgPheTrpAlaSer 228
QY 563 ATTTTGGATTACAGTCTCTCTCTGATATGATGATGATGATGATGATGATGATGATGATGAT 622
Db 229 MetMetGluAlaAlaSerCysProProAspTyrValProGluIlePheHisPheHis 248
QY 623 AGTACTACTGGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 682
Db 249 ThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGly 268
QY 683 AAGAAATATCTACTGTCTTATATATGTTGCTCTCAGTGCAGTGTGCTGCTGCTGCTGCTGCT 742
Db 269 LysLysHisProThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsn 288
QY 743 CGGTTTAAAGGACTCAAGTATTTCGCTTGAATACCTAGCTCTCTAGGTTATGCTGTT 802
Db 289 SerPheLysGlyIleLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaVal 308
QY 803 GTAGTGATGACACACAGGGGATCCTGTCCAGCGAGGCTTAAATTTGAAGGCGCTTTAAA 862
Db 309 ValValIleAspGlyArgGlySerCysGlnArgLysLeuArgPheGluGlyAlaLeuLys 328
QY 863 TATAAATGGGTCAAAATAGAAATGACATCGAGTGGAGGAGGATCCATATCTAGCTTCT 922
Db 329 AsnGlnMetGlyGlnValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGlu 348
QY 923 CGATATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 982
Db 349 LysTyrGlyPheIleAspLeuSerArgValAlaIleHisGlyIleTrpSerTyrGlyGlyPhe 368
```

```

QY 983 CTCTCCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTGTGCTATTTGCTGGGCC 1042
Db 369 LeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAla 388
QY 1043 CCAGTCACTCTGTGGATCTTCTATGATATGATATGATATGATATGATATGATATGATATG 1081
Db 389 ProValThrValTrpMetAlaTyrAspThrGlyTyrThr 401

RESULT 3
US-11-264-096-1593
; Sequence 1593, Application US/11264096
; Publication No. US20060084794A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546D1
; CURRENT APPLICATION NUMBER: US/11/264,096
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1593
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (183)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-11-264-096-1593

Alignment Scores:
Pred. No.: 8.11e-90 Length: 497
Score: 790.00 Matches: 176
Percent Similarity: 79.1% Conservative: 40
Best Local Similarity: 64.5% Mismatches: 57
Query Match: 50.6% Indels: 0
DB: 11 Gaps: 0

US-10-825-632-8 (1-1083) x US-11-264-096-1593 (1-497)
QY 263 GAAATAATGATTGATGCTGAAGAGGATCATAGTTCAGAGGCTGGTATAT 322
Db 129 GluValLeuAlaArgHisGlySerLysIleTrpValAsnGluThrLysLeuValTyr 148
QY 323 TTGGAAGCACCAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTACGTAAT 382
Db 149 PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAla 168
QY 383 CCGTGAGAGGTGACAGGCTGACTGACCGTGGCTACTCATCTTGTGTCATGTCAG 442
Db 169 AlaGlyGluIleValArgLeuThrThrProGlyPheSerHis**CysSerMetSerGln 188
QY 443 CACTGTGACTCTTTTATAAGTAAGTATAGTAACACAGAGAAATCCACACTGTGCTCCTT 502
Db 189 AsnPheAspMetPheValSerHisTyrSerValSerThrProCysValHisVal 208
QY 503 TACAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCACAAACAAAGGAATTTTGGGCCACC 562
Db 209 TyrLysLeuSerGlyProAspAspProLeuHisGlyGlnProArgPheTrpAlaSer 228
QY 563 ATTTTGGATTACAGTCTCTCTCTGATATGATGATGATGATGATGATGATGATGATGATGAT 622
Db 229 MetMetGluAlaAlaSerCysProProAspTyrValProGluIlePheHisPheHis 248
QY 623 AGTACTACTGGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 682
```



;; PRIOR APPLICATION NUMBER: US 60/128,705  
;; PRIOR FILING DATE: 1999-04-09  
;; PRIOR APPLICATION NUMBER: US 09/540,209  
;; PRIOR FILING DATE: 2000-04-04  
;; NUMBER OF SEQ ID NOS: 10444  
;; SEQ ID NO 6408  
;; LENGTH: 745  
;; TYPE: PRT  
;; ORGANISM: B.fragilis  
US-11-079-463-6408

Alignment Scores:  
Pred. No.: 6e-29 Length: 745  
Score: 372.00 Matches: 117  
Percent Similarity: 40.7% Conservative: 77  
Best Local Similarity: 24.5% Mismatches: 133  
Query Match: 19.2% Indels: 150  
DB: 11 Gaps: 17

US-10-825-632-8 (1-1083) x US-11-079-463-6408 (1-745)

```
QY 23 GGAGTCGCTACCTTTGTTCTCCAGAGAA-----TTTGATAGATATTCTGGCTATTGG 76
Db 202 GlyileProAspTrpValTyrgluGluGluPheGlyPheAsnArgAlaLeuGluPhe--- 220
QY 77 TGGTGTCCAAAGCTGAACACACTCCCGAGTGGTGGTAAATCTTAGAATTCATATGAA 136
Db 221 -----AsnAlaAspAsnThr-----MetLeuAlaTyrrVal 230
QY 137 GAAATGATGATCTGAGGTGGAAATTTATTCATGTTACATCCCTATGTTG-----GAA 190
Db 231 ArgPheAspGluSerGluValPro-----SerTyrrThrPheProLeuPheAlaGlyGlu 248
QY 191 ACAAGGAGGCGAGAT-----TCATCCGTTATCTTAA 223
Db 249 AlaProArgTyrrAspAlaLeuGlnAspTyrrProGlyGluTyrrTyrrTyrrProLys 268
QY 224 ACAGTCAGCAATCTTAAGTC-----ACTTTT----- 253
Db 269 AlaGlyTyrrProAsnSerLysValSerValHisThrPheAspIleLysSerLysValThr 288
QY 254 AGATGTCAGAAATATGATGATGCTGAGGAGGAGATC----- 292
Db 289 ArgGlnValLysLeuProIleAspAlaAspGlyTyrrIleProArgIleArgPheThrGln 308
QY 292 ----- 292
Db 309 AspProAsnLysLeuAlaIleMetThrLeuAsnArgHisGlnAsnArgPheAspMetTyrr 328
QY 292 ----- 292
Db 329 PheAlaAspProArgSerThrValCysLysLeuAlaLeuArgAspGluSerProTyrrTyrr 348
QY 293 -----ATAGTTGAT----- 301
Db 349 IleAsnGluAsnValPheAspAsnIleGlnPheTyrrProGluTyrrPheSerPheValSer 368
QY 301 ----- 301
Db 369 AspLysSerGlyTyrrProHisLeuTyrrTyrrTyrrSerMetAsnGlyAsnLeuLeuLysGln 388
QY 302 -----GAAGTCAGAGGCTGGTA----- 319
Db 389 ValThrSerGlyAsnTyrrGluValLysAsnPheIleGlyTyrrAsnProAspThrAsnGlu 408
QY 320 ----TATTTTGAAGGACCAAGACTCCCTTTAGAGCATACCTGTAGCTAGTACAGTTAC 376
Db 409 PheTyrrTyrrSerAsnGluGluSerProMetArgGlnAlaValTyrrLysIleAsp--- 427
QY 377 GTAATCCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTACATCTTCTGTGCATC 436
Db 428 ---ArgLysGlyLysMetLysLeuSerAsnGlnProGlyThrAsnSerProIlePhe 446
```

```
QY 437 AGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTG 496
Db 447 SerSerMetLysTyrrPheMetAsnLysPheThrSerLeuAspThrProMetLeuIle 466
QY 497 TCCCTTTACAAGCTATCAAGTCTGAGATGACCCAACTTGCAAAACAAAGAAATTTGG 556
Db 467 ThrLeu-----AsnAspAsnThrGlyLysValLeuLysThrLeu 479
QY 557 GCCACCAATTTGGATTGAGCAGGCTCTTCTCTGACTATATACCTCTCCA-----GAAATT 610
Db 480 ValThrAsnAspLysLeuLysGlnLysLeuAlaGluTyrrAlaIleProGlnLysGluPhe 499
QY 611 TTCTCTTTTGAAGTACTACTGATTTACATGTTATGGATGCTCTACAGCTCATGAT 670
Db 500 PheThrPheLysThrThrGluGlyValAspLeuAsnGlyTyrrMetMetLysProValAsn 519
QY 671 CTACAGCTGGAAGAAATATCTCTACTGTGCTGCTTATATATATGCTGCTCTCAGGTGCAG 730
Db 520 PheAspProAlaLysArgTyrrProValLeuMetPheGlnTyrrSerGlyProGlySerGln 539
QY 731 TTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTTCGGTTGTAATACCTTAGCTCTCTA 790
Db 540 GlnValLeuAspLysTrp---GlyIleSerTrp-----GluThrTyrrMetAlaSerLeu 556
QY 791 GGTATGTTGGTTAGTATAGATAGACAAACAGGGATCTCTGTCACCGAGGCTTAAATTTGAA 850
Db 557 GlyTyrrValValAlaCysValAspGlyArgGlyThrGlyGlyArgGlySerGluPheGln 576
QY 851 GCGCGCTTTAAATATAAATCGGTCAAATAGAAATTCACATGAGTGGAGGAGCTCCAA 910
Db 577 LysCysThrTyrrLeuAsnLeuGlyValLysGluAlaLysAspGlnValGluAlaLys 596
QY 911 TATCTAGCTTCPCGATATGATTTCATTGACTTAGTATGCTGTGGCATCCACGGCTGTCC 970
Db 597 TyrrLeuGlyGly---LeuProTyrrValAspLysGlyArgIleGlyIleTrpGlyTrpSer 615
QY 971 TATGAGGATACCTCTCCCTGATGCAATTAATGAGAGGTCAGATATCTTCAGGTTGCT 1030
Db 616 PheGlyGlyTyrrMetThrIleMetSerMetSerGluGlyThrProValPheLysAlaGly 635
QY 1031 ATTGCTGGGCGCCAGTCATCTCTGTCGATCTTCTATGATACAGGATACAG 1081
Db 636 ValAlaValAlaAlaProThrAspTrpLysTyrrTyrrAspThrValTyrrThr 652
RESULT 6
US-11-208-288-4
; Sequence 4, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-4
Alignment Scores:
Pred. No.: 3.03e-23 Length: 738
Score: 315.00 Matches: 120
Percent Similarity: 34.9% Conservative: 59
Best Local Similarity: 23.4% Mismatches: 146
Query Match: 16.3% Indels: 188
DB: 11 Gaps: 22
```



Db	206	AlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe-----	218
QY	119	CTTAGAATTCATATGAAGAAATGATGAATCTGAGGTGGAATTAATTCAT-----	169
Db	219	-----LeuAlaTyrAlaGlnPheAsnAspThrGluValProLeuIleGluTyrSerPhe	236
QY	170	GTTACATCCCTATGTTGGAACACAGGAGGCGAGATTCATTCGGTTATCCTAAACAGGT	229
Db	237	TyrSerAspGluSerLeuGlnTyrProLysThrValArgValProTyrProLysAlaGly	256
QY	230	ACAGCAATCTAAAGTCACCTTT-----	253
Db	257	AlaValAsnProThrValLysPhePheValValAsnThrAspSerLeuSerSerValThr	276
QY	254	AAAGTGTGAGAAATAATCATGATGTGAAGGAAGGATCATGTTGATGAGTCAAGG	313
Db	277	AsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIleGly-----	292
QY	314	CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCCTGTACGTAGTCAGT	373
Db	293	-----AspHisTyrLeuCysAspValThr	300
QY	374	TACGTAATCTGGAGAGGTGACA-----AGGCTGACTGACCGTGGCTAC	418
Db	301	TrpAlaThrGlnGluArgIleSerLeuGlnTrpLeuArgArgIleGlnAsnTyrSerVal	320
QY	419	TCACATTTCTGTC-----TGCATC-----AGTCAG	442
Db	321	MetAspIleCysAspTyrAspGluSerSerGlyArgTrpAsnCysLeuValAlaArgGln	340
QY	443	CACGTGACTTC-----TTTATAGTAAGTATAGTAACAGAGAATCCACAC	490
Db	341	HisIleGluMetSerThrThrGlyTrpValGlyArgPhe---ArgProSerGluProHis	359
QY	491	TGTGTG-----TCCCTTTACAGCTCAAGTCTCTGAAGAT-----	526
Db	360	PheThrLeuAspGlyAsnSerPheTyrIlysIleSerAsnGluGluGlyTyrArgHis	379
QY	527	-----GACCAACTGCAAAACAAAGGAATTTGGGCC	559
Db	380	IleCysTyrPheGlnIleAspIysLysAspCysThrPheIleThrIysGlyThrTrpGlu	399
QY	560	ACCAT-----	565
Db	400	ValIleGlyIleGluAlaLeuThrSerAspTyrLeuTyrIleSerAsnGluTyrLys	419
QY	565	-----	565
Db	420	GlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuSerAspTyrThrLysValThr	439
QY	565	-----	565
Db	440	CysLeuSerCysGluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLys	459
QY	566	-----TTGGATTGAGGAGTCTCT-----CTTCTGACTACTACT-----	598
Db	460	GluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHis	479
QY	598	-----	598
Db	480	SerSerValAsnAspLysGlyLeuArgValLeuGluAspAsnSerAlaLeuAspLysMet	499
QY	599	-----CTCCAGAAATTTCTCTTT-----GAAAGTACTACTGGA	634
Db	500	LeuGlnAsnValGlnMetProSerLysLysLeuAspPheIleIleLeuAsnGluThrLys	519
QY	635	TTTACATGTTGAGTCTCTACAAAGCTCTCATGATCTACAGCTTGGAAAGAAATATCCT	694
Db	520	Phe---TrpTyrGlnMetIleLeuProProHis---PheAspLysSerLysLysTyrPro	537
QY	695	ACTGTGTTGATATATGTTGCTCTCAGGTGCGATTAATCGGTTAAAGGA	754
Db	538	LeuLeuAspValTyrAlaGlyProCysSerGln-----LysAla	551
QY	755	GTCAGTATTTCCTCGTTGAAT-----ACCTAGCCTCTCTAGGTTATGTGTTGTA	805
Db	552	AspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleVal	571
QY	806	GTG---APAGACAACAGGGATCTGTCCAGGAGGCTTAATTTTGAAGCGCCTTAAA	862
Db	572	AlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsn	591
QY	863	TATAAATGGGTCAAAATAGAAATGACATCAGTGGAGGACTCCCAATATCTAGCTTCT	922
Db	592	ArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe---Ser	610
QY	923	CGATATGATTTTCATTTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATAC	982
Db	611	LysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGlyTyr	630
QY	983	CTCTCCCTGATGGCATTATTCAGAGTCCAGATATCTTCAGGTTGCTATTGCTGGGCC	1042
Db	631	ValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAla	650
QY	1043	CCAGTCACTCTGTGATCTTCTATCATACAGGATACACG	1081
Db	651	ProValSerArgTrpGluTyrTyrAspSerValTyrThr	663
RESULT 8			
US-10-522-789-2			
; Sequence 2, Application US/10522789			
; Publication No. US20050260732A1			
; GENERAL INFORMATION:			
; APPLICANT: TANABE SEIYAKU CO., LTD.			
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV			
; FILE REFERENCE: 03-039-PCT			
; CURRENT APPLICATION NUMBER: US/10/522,789			
; CURRENT FILING DATE: 2005-01-28			
; PRIOR APPLICATION NUMBER: US 60/398,761			
; PRIOR FILING DATE: 2002-07-29			
; NUMBER OF SEQ ID NOS: 2			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 2			
; LENGTH: 766			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-522-789-2			
Alignment Scores:			
Pred. No.:	3,05e-23	Length:	766
Score:	315.00	Matches:	120
Percent Similarity:	34.9%	Conservative:	59
Best Local Similarity:	23.4%	Mismatches:	146
Query Match:	16.3%	Indels:	188
DB:	9	Gaps:	22
US-10-825-632-8 (1-1083) x US-10-522-789-2 (1-766)			
QY	2	GAAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA-----TTTGAT	58
Db	190	LysGluAspIleLeuTyrAsnGlyIleThrAspTrpValTyrGluGluGluValPheSer	209
QY	59	AGATATTCTGGTATTGTTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGAATAATT	118
Db	210	AlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe-----	222
QY	119	CTTAGAATTCATATGAAGAAATGATGAATCTGAGGTGGAATTAATTCAT-----	169
Db	223	-----LeuAlaTyrAlaGlnPheAsnAspThrGluValProLeuIleGluTyrSerPhe	240
QY	170	GTTACATCCCTATGTTGGAACACAGGAGGCGAGATTCATTCGGTTATCCTAAACAGGT	229
Db	241	TyrSerAspGluSerLeuGlnTyrProLysThrValArgValProTyrProLysAlaGly	260
QY	230	ACAGCAATCTTAAAGTCACCTTT-----	253

Db 261 AlaValAsnProThrValLysPheValValAsnThrAspSerLeuSerSerValThr 280  
QY 254 AAGATGTCAGAAATAATGATGCTGAAGGAGGATCATAGTTGATGAAGTCAGAGG 313  
Db 281 AsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIleGly 296  
QY 314 CTGATATATTTGAAGGCACCAAGATCCCTTTAGAGCATCACTGATAGTAGTCAGT 373  
Db 297 -----AspHisTyrLeuCysAspValThr 304  
QY 374 TAGCTAAATCCTCGAGAGTGACA-----AGCTGACTGACCGTGGCTAC 418  
Db 305 TrpAlaThrGlnGluArgIleSerLeuGlnTrpLeuArgArgIleGlnAsnTyrSerVal 324  
QY 419 TCACATCTCTGC-----TCATC-----AGTCAG 442  
Db 325 MetAspIleCysAspTyrAspGluSerSerGlyArgTrpAsnCysLeuValAlaAaGln 344  
QY 443 CACTGTGACTTC-----TTTATAAGTAAATAGTAAACCAAGAAATCCACAC 490  
Db 345 HisIleGluMetSerThrThrGlyTrpValGlyArgPhe---ArgProSerGluProHis 363  
QY 491 TGTGTG-----TCCCTTTACAACTCAAGTCTCTGAAGAT-----526  
Db 364 PheThrLeuAspGlyAsnSerPheTyrLysIleIleSerAsnGluGluGlyTyrArgHis 383  
QY 527 -----CACCAACTTGCAAAACAAAGAAATTTTGGGCC 559  
Db 384 IleCysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGlu 403  
QY 560 ACCATT-----565  
Db 404 ValIleGlyIleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLys 423  
QY 565 -----565  
Db 424 GlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuSerAspTyrThrLysValThr 443  
QY 565 -----565  
Db 444 CysLeuSerCysGluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLys 463  
QY 566 -----TTGATTCTAGCAGGTCCT---CTTCTCTACTACT---598  
Db 464 GluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHis 483  
QY 598 -----598  
Db 484 SerSerValAsnAspLysGlyLeuArgValLeuGluAspAsnSerAlaLeuAspLysMet 503  
QY 599 -----CCTCCAGAAATTTCTCTTTT-----GAAAGTACTACTGGA 634  
Db 504 LeuGlnAsnValGlnMetProSerLysLysLeuAspPheIleLeuAsnGluThrLys 523  
QY 635 TTTACATTGTATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCCT 694  
Db 524 Phe---TrpTyrGlnMetIleLeuProProHis---PheAspLysSerLysLysTyrPro 541  
QY 695 ACTGTCTGTTCATATATGTTGCTCTCAGGTGCAAGTGTGTAATAATCGGTTTAAAGGA 754  
Db 542 LeuLeuLeuAspValTyrAlaGlyProCysSerGln-----LysAla 555  
QY 755 GTCAAGTATTTCGCTTGAAT-----ACCCTAGGCTCTCTAGGTTATGTGGTTGTA 805  
Db 556 AspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleVal 575  
QY 806 GTG---ATAGACAAACAGGGATCCTGTCAACCGAGGCTTAAATTTGAAGGCGCTTTAAA 862  
Db 576 AlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsn 595  
QY 863 TATAAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAGGACTCCCAATATCTAGCTTCT 922  
Db 596 ArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe---Ser 614

QY 923 CGATATGATTTTCATTGACTTAGATCGTGTGGCATCCACGGCTGCTCTATGGAGATAC 982  
Db 615 LysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTyrSerTyrGlyTyr 634  
QY 983 CTCCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGTTGCTATTGCTGGGCC 1042  
Db 635 ValThrSerMetValLeuGluSerGlySerGlyValPheLysCysGlyIleAlaValAla 654  
QY 1043 CCAGTCACTCTGTGGATCTCTATGATACAGATACAG 1081  
Db 655 ProValSerArgTrpGluTyrTyrAspSerValTyrThr 657

## RESULT 9

US-10-501-035-234

; Sequence 234, Application US/10501035

; Publication No. US20060046249A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING

; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASES

; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS

; FILE REFERENCE: D0185 PCT

; CURRENT APPLICATION NUMBER: US/10/501,035

; PRIOR FILING DATE: 2004-07-09

; PRIOR APPLICATION NUMBER: US 60/350,061

; NUMBER OF SEQ ID NOS: 795

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 234

; LENGTH: 766

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-501-035-234

## Alignment Scores:

Pred. No.:	3,05e-23	Length:	766
Score:	315.00	Matches:	120
Percent Similarity:	34.9%	Conservative:	59
Best Local Similarity:	23.4%	Mismatches:	146
Query Match:	16.3%	Indels:	188
DB:	9	Gaps:	22

US-10-825-632-8 (1-1083) x US-10-501-035-234 (1-766)

QY 2 GAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTCTCCCAAGAGAA---TTTGAT 58

Db 190 LysGluAspIleIleTyrAsnGlyIleThrAspTrpValTyrGluGluValPheSer 209

QY 59 AGATATTCTGCTATTGCTGTGTCCAAAGCTGAAACAACTCCAGTGGTGTAAATTT 118

Db 210 AlaTyrSerAlaLeuTyrTrpSerProAsnGlyThrPhe-----222

QY 119 CTAGAATTCTATATGAAGAAATGATGATCTGAGTGGAAATTTTTCAT-----169

Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluValProLeuIleGluTyrSerPhe 240

QY 170 GTTACATCCCCTATGTTGGAAACAAGGAGGCGAGATTCATTCGTTATCCTAAACAGGT 229

Db 241 TyrSerAspGluSerLeuGlnTyrProLysThrValArgValProTyrProLysAlaGly 260

QY 230 ACAGCAATCCTAAAGTCACTTT-----253

Db 261 AlaValAsnProThrValLysPhePheValValAsnThrAspSerLeuSerValThr 280

QY 254 AAGATGTCAGAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 313

Db 281 AsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIleGly-----296

QY 314 CTGATATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCTCAGT 373

Db 297 -----AspHisTyrLeuCysAspValThr 304



```
QY 374 TAGCTAAATCCTGGAGAGGTGACA-----AGCTGACTGACCGCTGGCTAC 418
Db 305 TtpAlaThrGlnGluArgIleSerLeuGlnTrpLeuArgArgIleGlnAsnTrpSerVal 324
QY 419 TCACATTCCTTGC-----TGCATC-----AGTCAG 442
Db 325 MetAspIleCysAspTyrAspGluSerSerGlyArgTrpAsnCysLeuValAlaAaArgGln 344
QY 443 CACTGTGACTTC-----TTTATAAGTAAATAGTATAGTAAACAGAAAGAAATCCACAC 490
Db 345 HisIleGluMetSerThrThrGlyTrpValGlyArgPhe---ArgProSerGluProHis 363
QY 491 TGTGTG-----TCCCTTTACAAAGTATCAAGTCTGAGAT-----526
Db 364 PheThrLeuAspGlyAsnSerPheTyrLysIleSerAsnGluGlyTyrArgHis 383
QY 527 -----GACCAACTTGCMAAACCAAGGAATTTGGGCC 559
Db 384 IleCysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGlu 403
QY 560 ACCATT-----565
Db 404 ValIleGlyIleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsnGluTyrLys 423
QY 565 -----565
Db 424 GlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyrThrLysValThr 443
QY 565 -----565
Db 444 CysLeuSerCysGluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLys 463
QY 566 -----TTGGAATTCACAGAGTCTCT-----598
Db 464 GluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHis 483
QY 598 -----598
Db 484 SerSerValAsnAspLysGlyLeuArgValLeuGluAspAsnSerAlaLeuAspLysMet 503
QY 599 -----CCTCCAGAAATTTTCTCTTT-----GAAAGTACTACTGGA 634
Db 504 LeuGlnAsnValGlnMetProSerLysLysLeuAspPheIleIleLeuAsnGluThrLys 523
QY 635 TTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAGAAATATCCT 694
Db 524 Phe---TrpTyrGlnMetIleLeuProHis---PheAspLysSerLysLysTyrPro 541
QY 695 ACTGTGCTGTATATATGCTGCTCAGGTGCGAGTTGGTGAATAATCGGTTTAAAGGA 754
Db 542 LeuLeuLeuAspValTyrAlaGlyProCysSerGln-----LysAla 555
QY 755 GTCAAGTATTCGGTTGAAT-----ACCTAGCCTCTCTAGGTTATGTGGTTGTA 805
Db 556 AspThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleVal 575
QY 806 GTG---ATAGACAACAGGGATCCTGCTCACCGAGGCTTAAATTTGAAGGCGCTTTAAA 862
Db 576 AlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsn 595
QY 863 TATAAATGGTCAAAATAGAAATGACGATCAGGTGGAGGACTCCAATATCTAGCTTCT 922
Db 596 ArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAaArgGlnPhe---Ser 614
QY 923 CGATATGATTTCAATGACTAGATCGTGTGGGCATCCACGGCTGCTATGAGGATAC 982
Db 615 LysMetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGlyTyr 634
QY 983 CTCTCCCTGATGCATTAATAGCAGAGTTCAGATATCTTCAGGTTGCTATTGCTGGGCC 1042
Db 635 ValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAla 654
QY 1043 CCAGTCACTGTGGATCTTCTATGATACAGGATACAG 1081
```

```
Db 655 ProValSerArgTrpGluTyrTyrAspSerValTyrThr 667
RESULT 10
US-11-208-288-2
; Sequence 2, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-2
Alignment Scores:
Pred. No.: 3,05e-23 Length: 766
Score: 315.00 Matches: 120
Percent Similarity: 34.9% Conservative: 59
Best Local Similarity: 23.4% Mismatches: 146
Query Match: 16.3% Indels: 188
DB: 11 Gaps: 22
US-10-825-632-8 (1-1083) x US-11-208-288-2 (1-766)
QY 2 GAAGAAGATCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAGAAGAA---TTTGAT 58
Db 190 LysGluAspIleIleTyrAsnGlyIleThrAspTrpValTyrGluGluValPheSer 209
QY 59 AGATATTCTGGCTATTGGTGTCTCCAAAAGCTCCCAACCACTCCAGTGGTGGTAAAT 118
Db 210 AlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe-----222
QY 119 CTAGAATTCTATGAAGAAATATGATGAATCTGAGTGGAAATTTATTCAT-----169
Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluValProLeuIleGlyTyrSerPhe 240
QY 170 GTTACATCCCTATGTTGGAAACAGGAGCGAGATTCATTCGGTTATCCTTAAACAGGT 229
Db 241 TyrSerAspGluSerLeuGlnTyrProLysThrValArgValProTyrProLysAlaGly 260
QY 230 ACAGCAATCTTAAAGTCACCTTT-----253
Db 261 AlaValAsnProThrValLysPhePheValValAsnThrAspSerLeuSerSerValThr 280
QY 254 AAGATGTCAGAAATAATGATGATGCTGAAGGAGAGATCATAGTTGATGAAGTCAGAGG 313
Db 281 AsnAlaThrSerIleGlnIleThrAlaProAlaSerMetLeuIleGly-----296
QY 314 CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCCTCTGTCGTAGTCAGT 373
Db 297 -----AspHisTyrLeuCysAspValThr 304
QY 374 TACGTAAATCTCGAGAGGTGACA-----AGCTGACTGACCGTGGCTAC 418
Db 305 TrpAlaThrGlnGluArgIleSerLeuGlnTrpLeuArgArgIleGlnAsnTrpSerVal 324
QY 419 TCACATTCCTTGC-----TGCATC-----AGTCAG 442
Db 325 MetAspIleCysAspTyrAspGluSerSerGlyArgTrpAsnCysLeuValAlaAaArgGln 344
QY 443 CACTGTGACTTC-----TTTATAAGTAAATAGTATAGTAAACAGAAAGAAATCCACAC 490
Db 345 HisIleGluMetSerThrThrGlyTrpValGlyArgPhe---ArgProSerGluProHis 363
```

```
QY 491 TGTGTG-----TCCCTTTACAAGCTATCAAGTCCTCGAAGAT----- 526
Db 364 PheThrLeuAspGlyAenSerPheTyrLysIleSerAenGluGluGlyTyrArgHis 383
QY 527 -----GACCCAATCTGCAAAACAAGAGATTTTGGGCC 559
Db 384 IleCysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrPglu 403
QY 560 ACCATT----- 565
Db 404 ValIleGlyLeuAlaLeuThrSerAspTyrLeuTyrTyrIleSerAenGluTyrLys 423
QY 565 ----- 565
Db 424 GlyMetProGlyGlyArgAenLeuTyrLysIleGlnLeuIleAspTyrThrLysValThr 443
QY 565 ----- 565
Db 444 CysLeuSerCysGluLeuAenProGluArgCysGlnTyrTyrIleSerValSerPheSerLys 463
QY 566 -----TTGATTTCAGCAGGTCTCT----- 598
Db 464 GluAlaLysTyrTyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHis 483
QY 598 ----- 598
Db 484 SerSerValAenAspLysGlyLeuArgValLeuGluAenSerAlaLeuAspLysMet 503
QY 599 -----CCTCCAGAAATTTCTCTTTT-----GAAAGTACTACTGGA 634
Db 504 LeuGlnAenValGlnMetProSerLysLysLeuAspPheIleIleLeuAenGluThrLys 523
QY 635 TTTACATTGTTGGGATGCTTACAAAGCTCATGATCTACAGCTCGAAGAAATATCTCT 694
Db 524 Phe---TrpTyrGlnMetIleLeuProProHis---PheAspLysSerLysTyrPro 541
QY 695 ACTGTGCTGTTCAATATGTTGGTCTCAGTCAGTGTGTAATATCGGTTTAAAGGA 754
Db 542 LeuLeuLeuAspValTyrAlaGlyProCysSerGln-----LysAla 555
QY 755 GTCAAGTATTTCCGTTGAAT-----ACCTAGCTCTCTAGGTATGCTGTTGTA 805
Db 556 AspThrValPheArgLeuAenTrpAlaThrTyrIleAenSerThrGluAenIleVal 575
QY 806 GTG---ATAGACAACAGGGATCTGTCCAGGAGGCTTAATTTGAAGCGCCCTTTAAA 862
Db 576 AlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAen 595
QY 863 TATAAATGGCTCAATAGAAATGACATCAGTGAAGGACTCCAAATATCTAGCTTCT 922
Db 596 ArgArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe---Ser 614
QY 923 CGATATGATTTCATTGACTTAGATCGTGTGGGCATCCAGCGTGTCTCTATGGAGATAC 982
Db 615 LysMetGlyPheValAspAenLysArgIleAlaIleTrpGlyTrpSerTyrGlyGlyTyr 634
QY 983 CTCCTCCCTGATGGCATTAAATGACAGAGTTCAGATATCTTCAGGGTTGCTATTGGGGGCC 1042
Db 635 ValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAla 654
QY 1043 CCATCTACTCTGTGGATCTTCTATGATACAGATACACG 1081
Db 655 ProValSerArgTrpGluTyrTyrAspSerValTyrThr 667
```

## RESULT 11

```
US-11-208-288-6
; Sequence 6, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
```

```
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; PRIOR FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-208-288-6
```

## Alignment Scores:

Pred. No.:	1,21e-22	Length:	760
Score:	309.00	Matches:	128
Percent Similarity:	37.1%	Conservative:	57
Best Local Similarity:	25.7%	Mismatches:	154
Query Match:	15.9%	Indels:	160
DB:	11	Gaps:	23

US-10-825-632-8 (1-1083) x US-11-208-288-6 (1-760)

```
QY 2 GAAGAGATGCCAGATCAGCTCGTACTCTTGTCTCCAGAGAA---TTTGAT 58
Db 184 GluGlnAenValIleTyrAenGlyIleThrAspTrpValTyrGluGluValPheGly 203
QY 59 AGATATCTCGCTATTGGTGGTGTCCAAA-----GCTGAACAACACT 100
Db 204 AlatySerAlaLeuTrpTrpSerProAenAenThrPheLeuAlaTyrAlaGlnPheAen 223
QY 101 CCAGTGGTGGTAAATCTTAGAATCTATATGAAGAAATATGATGATCTGAGGTGAA 160
Db 224 AspThrGlyValProLeuIleGlnTyrSerPhe---TyrSerAspGluSer--- 239
QY 161 ATTATTCATGTTTACATCCCTCTATGTGTGAACAAGGAGGCAGATTTCCTGTTATCCT 220
Db 240 -----LeuGlnTyrProLysThrValTyrPheProTyrPro 251
QY 221 AAAACAGGTACAGCAAACTCTAAAGTCACTTT----- 253
Db 252 LysAlaGlyAlaValAenProThrValLysPhePheIleValAenIleAspSerLeuSer 271
QY 253 ----- 253
Db 272 SerSerSerAlaAlaProIleGlnIleProAlaProAlaSerValAlaArgGlyAsp 291
QY 254 ---AAGATGTCAGAAATAATGATTGCTGCTGAAGGAGGATCATAGTTGATGAGTCAGA 310
Db 292 HisTyrLeuCysAspValValTyrAlaThrGluGluArgIleSerLeuGlnTrpLeuArg 311
QY 311 AGGCTGGTATATTGTGAAGC-----ACCAAA 337
Db 312 ArgIleGlnAenTyrSerValMetAlaIleCysAspTyrAspLysIleAenLeuThrTrp 331
QY 338 GACTCCCTCTTAGAG---CATCACCCTGACGTA-----GTCAGTTACGTA 379
Db 332 AenCysProSerGluGlnGlnHisValGluMetSerThrThrGlyTrpValGlyArgPhe 351
QY 380 AATCTGAGAGGTGACAAAGCTGACGTGAC----- 409
Db 352 ArgProAlaGluProHisPheThrSerAspGlySerSerPheTyrLysIleIleSerAsp 371
QY 410 ---CGTGGTACTCACATTTCTCTGTCATC-----AGTCAGCAGCTGTGACTTC--- 454
Db 372 LysAspGlyTyrLysHisIleCysHisPheProLysAspLysLysAspCysThrPheIle 391
QY 455 -----TTTATAAGTAAG 466
Db 392 ThrLysGlyAlaTrpGluValIleSerIleGluAlaLeuThrSerAspTyrLeuTyrTrp 411
QY 467 TATAGTAACCAAG-----AAGAAATCCACACTGTGTCTCCCTTTTACAAG----- 508
```

Db 412 IleSerAsnGlnTyrLysGluMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuThr 431  
Qy 509 -----CTATCAAGTCTGAGATGACCCAACTTGGAAA----- 541  
Db 432 AspHisThrAsnValLysCysLeuSerCysAspLeuAsnProGluAtrGlyGlnTyrTyr 451  
Qy 542 -----ACAAAGGAATTTTGGGCCACCATTTTGGATTCACAGAGTCTCT--CTT 586  
Db 452 AlavalSerPheSerLysGluAlaLysTyrTyrGlnLeuGlyCysTrpGlyProGlyLeu 471  
Qy 587 CTGACTATACT----- 598  
Db 472 ProLeuTyrThrLeuHisArgSerThrAspHisLysGluLeuArgValLeuGluAspAsn 491  
Qy 599 -----CTCCAGAAATTTTCTCTTTT----- 619  
Db 492 SerAlaLeuAspArgMetLeuGlnAspValGlnMetProSerLysLysLeuAspPheIle 511  
Qy 620 ---GAAAGTACTACTGATTTACATTGTGGATGCTCTACAGCCCTCATGATGATCTACAG 676  
Db 512 ValLeuAsnGlnThrArgPhe--TrpTyrGlnMetIleLeuProProHis--PheAsp 529  
Qy 677 CTGGGAAGAATATCTACTGCTGTTCATATATGTTGCTGCTCCTCAGGTGCAGTTCGTG 736  
Db 530 LysSerLysTyrProLeuLeuAspValTyrAlaGlyProCysSerGln----- 547  
Qy 737 AATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAAT-----ACCTAGCCCTCT 787  
Db 548 -----LysAlaAspAlaSerPheArgLeuAsnTrpAlaThrTyrLeuAlaSer 563  
Qy 788 CTAGGTATGTTGTTGTAGT---ATAGACAACAGGGATCCTGTCACCGAGGCTTAAA 844  
Db 564 ThrGluAsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLys 583  
Qy 845 TTTGAAGGCGCTTTAAATATAATAATGGTCAAAATAGAAATGACGATCAGTGGGAAGGA 904  
Db 584 IleMetHisAlaIleAsnArgLeuGlyThrLeuGluValGluAspGlnIleGluAla 603  
Qy 905 CTCCAATATCTAGTCTTCGATATGATTTCAATTGATTTAGATCGTGGGATCCACGGC 964  
Db 604 AlaArgGlnPheVal---LysMetGlyPheValAspSerLysArgValAlaIleTrpGly 622  
Qy 965 TGCTCTATGAGGATACCTCTCCCTGATGGATTAATGACAGAGTCAGATATCTTCAGG 1024  
Db 623 TrpSerTyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLys 642  
Qy 1025 GTTGTCTATGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGC 1081  
Db 643 CysGlyIleAlaValAlaProValSerArgTrpGluTyrTyrAspSerValTyrThr 661

## RESULT 12

US-11-186-284-55  
; Sequence 55, Application US/11186284  
; Publication No. US20050266493A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: BURGART, Lawrence J.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; FILE REFERENCE: MEMO1-029P2RNM  
; CURRENT APPLICATION NUMBER: US/11/186,284  
; CURRENT FILING DATE: 2005-07-21  
; PRIOR APPLICATION NUMBER: US/10/301,822  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339,971  
; PRIOR FILING DATE: 2001-12-10

; PRIOR APPLICATION NUMBER: US 60/361,978  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/381,988  
; PRIOR FILING DATE: 2002-05-20  
; NUMBER OF SEQ ID NOS: 228  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 55  
; LENGTH: 760  
; TYPE: PRF  
; ORGANISM: Homo Sapiens  
; US-11-186-284-55  
Alignment Scores:  
Pred. No.: 1,94e-17 Length: 760  
Score: 257.00 Matches: 110  
Percent Similarity: 36.8% Conservative: 68  
Best Local Similarity: 22.7% Mismatches: 158  
Query Match: 13.3% Indels: 148  
DB: 11 Gaps: 20  
US-10-825-632-8 (1-1083) x US-11-186-284-55 (1-760)  
Qy 23 GGAGTCGCTACCTTTGTTCTCCAGAGAAGATTT-----GATAGATATTCTGGCTATTGG 76  
Db 195 GlyIleProAspTrpValTyrGluGluMetLeuProThrLysTyrAla---LeuTrp 213  
Qy 77 TGGTGTCCAAAGCTGAAACAACACTCCAGTGGTGGTAAATAATCTTAGAATTTCTATATGAA 136  
Db 214 TrpSerPro-----AsnGlyLysPheLeu-----AlaTyrAla 224  
Qy 137 GAAATGATGATGATCGAGTGGAAATATTTCATGTTTACATCCCTATGTTGGAACA---- 193  
Db 225 GluPheAsnAspLysAspIleProValIleAlaTyrSerTyrTyrGlyAspGluGlnTyr 244  
Qy 194 AGGAGGCGAGATTCATTCCGTTATCTCTAAACAGGTACAGCAATCTCTAAAGTCACCTTT 253  
Db 245 ProArgThrIleAsnIleProTyrProLysAlaGlyAlaLysAsnProValValArgIle 264  
Qy 254 AAGATG-----TCAGAAATAATGATGATGCT 280  
Db 265 PheIleIleAspThrThrTyrProAlaTyrValGlyProGlnGluValProValProAla 284  
Qy 281 -----GAAGGAGGATC 292  
Db 285 MetIleAlaSerSerAspTyrTyrPheSerTrpLeuThrTrpValThrAspGluArgVal 304  
Qy 293 ATAGTTGATGAAGTCAGAGCGCTGATATATTTTGA-- 328  
Db 305 CysLeuGlnTrpLeuLysArgValGlnAsnValSerValLeuSerIleCysAspPheArg 324  
Qy 329 -----GGCACCACAAAGACTCCCT---TTAGACATCACCCTGTAC----- 364  
Db 325 GluAspTrpGlnTrpTrpAspCysProLysThrGlnGluHisIleGluGluSerArgThr 344  
Qy 365 -----GTAGTCAGTTTACGTAAATCCTGGAGAG 391  
Db 345 GlyTrpAlaGlyGlyPhePheValSerArgProValPheSerTyrAspAlaIleSerTyr 364  
Qy 392 GTGCAAGGCTGATGACCGT---GGCTACTCATCTTCTGTGTCATC----- 436  
Db 365 TyrIlystlePheSerAspLysAspGlyTyrLysHisIleHisTyrIleLysAspThrVal 384  
Qy 437 -----ACT 439  
Db 395 GluAsnAlaIleGlnIleThrSerGlyLysTyrGluAlaIleAsnIlePheArgValThr 404  
Qy 440 CAGCAGCTGTGACTTCTTTTATAAGT-----AAGTATAGTACCAAGAGATCCA 487  
Db 405 GlnAspSerLeuPheTyrSerSerAsnGluPheGluGluTyrProGlyArgArgAsnIle 424  
Qy 488 CACTGTGTGCTCCCTTTACAAAGCTATCAGTCTGAGAGATGACCCCACTTGTCAAAACAAAG 547  
Db 425 TyrArgIleSerIleGlySerTyrProProSerLysCysValThrCysHisLeuArg 444

```
QY 548 GAA-----TTTGGCCACCATTGTTGGATTGACGA----- 577
Db 445 LysGluArgCysGlnTyrThrAlaSerPheSerAspTyrAlaLysTyrThrAlaLeu 464
QY 578 -----GGTCCT---CTTCCTGACTATCT----- 598
Db 465 ValCysTyrGlyProGlyIleProIleSerThrLeuHisAspGlyArgThrAspGlnGlu 484
QY 599 ----- 601
Db 485 IleLysIleLeuGluGluAenLysGluLeuGluAenAlaLeuLysAenIleGlnLeuPro 504
QY 602 CCAGAAATTTCTCTTTTGAAGTACTACTCGATTACATTGTATGGGATGCTCTCAAG 661
Db 505 LysGluGluIleLysLysLeuGluValAspGluIleThrLeuThrLysMetIleLeu 524
QY 662 CCTCATGATCAGCCTGGAAAGAAATATCTACTGTGCTGTGTTTCATATATGTGTCT 721
Db 525 ProProGlnPheAspArgSerLysLysTyrProLeuLeuIleGlnValTyrGlyGlyPro 544
QY 722 CAGGTGCAGTTGGTGAATAATCGGTTTAAAGGATCAAGTATTTCGCTTGAATACCTTA 781
Db 545 CysSerGlnSerValArgSer-----ValPheAlaValAenTrpIle 558
QY 782 GCCTCTCTA-----GGTTATGCTGTGATGATGATGATGATGATGATGATGATGAT 829
Db 559 SerTyrLeuAlaSerLysGluGlyMetValIleAlaLeuValAspGlyArgGlyThrAla 578
QY 830 CACCCAGGGCTTAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAATTCAC 889
Db 579 PheGlnGlyAspLysLeuLeuTyrAlaValTyrArgLysLeuGlyValTyrGluValGlu 598
QY 890 GATCAGGTGGAGGATCCCAATATCTAGCTTCTCGATATGATTTCATTGACATTAGATCGT 949
Db 599 AspGlnIleThrAlaValArgLysPheIle---GluMetGlyPheIleAspGluLysArg 617
QY 950 GTGGGATCCAGCGCTGTCTCTATGAGGATACCTCTCCCTGATGATGATGATGATGATGAT 1009
Db 618 IleAlaIleThrAlaValArgLysPheIle---GluMetGlyPheIleAspGluLysArg 617
QY 1010 TCAGATATCTTCAGGTTGTATTGCTGGGCGCCAGTCACTCTGCTGGATCTCTCATGAT 1069
Db 638 ThrGlyLeuPheLysCysGlyIleAlaValAlaProValSerSerTrpGluTyrThrAla 657
QY 1070 ACAGGATACAG 1081
Db 658 SerValTyrThr 661

RESULT 13
; Sequence 7504, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FRAGILIS
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/11/079,463
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7504
; LENGTH: 624
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-7504

Alignment Scores: 1.19e-11 Length: 624
Pred. No.: 624
```

```
Score: 199.00 Matches: 93
Percent Similarity: 37.9% Conservative: 50
Best Local Similarity: 24.7% Mismatches: 140
Query Match: 10.3% Indels: 94
DB: 11 Gaps: 18
US-10-825-632-8 (1-1083) x US-11-079-463-7504 (1-624)
QY 56 GATAGATATTCTGGCTATTGGTGTCTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAA 115
Db 189 AspArgPheLeuGly-----GlyAlaSerTyrSerProAspGlyLysGln 203
QY 116 ATTCCTTGAATTTCTATATGAAGAA----- 139
Db 204 LeuLeuLeuThrAlaSerProGluAlaPheAspGlyIleGlyLysAsnCysGlyAsnHis 223
QY 140 -----AATCAT---GAATCTGAGGTGGAATATTCATGTTTACATCCCTCATGTTG 187
Db 224 ProIleAlaAsnAspPheAspSerGlnAlaPheIleMetAsp-----Leu 238
QY 188 GAAACAAGAGGAGGAGATTCATTCGTTATCTCTAAACAGGTACAGCAATCTTAAAGTC 247
Db 239 AlaThrArgLysIleAsp-----ProIleThrLysGluPheAsnProSerVal 254
QY 248 ACTTTTAAGATGTCAGAAATAATGATGCTGAGGAAGAGATCATAGTT----- 298
Db 255 AsnPheLeuGlnTrpAsn-----LysGlyAspGlyCysIleTyrPheSerThrAsn 271
QY 299 GATGAGTCTCAGAGGCTGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCAT--- 355
Db 272 AspGluAspCysArgAsnIleTyrArgTyrSerProLysAspArgLysPheGluLysLeu 291
QY 356 -----CACCTGTAAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 394
Db 292 AsnLeuGluThrAspValThrSerAlaPheAlaMetSerGluAsnAsnPro----- 308
QY 395 ACAAGCTGACTGACCGGTCTACTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 454
Db 309 -----SerLeuAlaAlaTyrIleGlyGlnGlyCys----- 318
QY 455 TTTTATAAGTATATAGTAAACAGAGAAATTTTGGGCCACCATTTTGGATTCA 514
Db 319 -----TyrAsnAlaGlyValAlaTyrValTyrAspLeuLys 330
QY 515 AGTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCA 574
Db 331 Lys-----LysThrSerArgLeu-----IleAlaAspPro 340
QY 575 GCAGGTCTCT-----CTTCCTGACTATATCTCTCCAGAAATTTTCTCT 616
Db 341 MetLysProThrLeuGluLysIleGluLeuGlyGluMetLysPro-----TrpAsn 357
QY 617 TTTGAAAGTACTACTGGATTTTACATTGTATGGGATGCTCTACAGGCTCATGATCTACAG 676
Db 358 PheThrAlaSerAspGlyThrGluIleLysGlyMetCysLeuProProSerPheAsp 377
QY 677 CCTGGAAGAAATATCTCTGCTGCTGCTCATATATGTTGCT-----CCTCAGGTGCAG 730
Db 378 ProAsnLysLysTyrProLeuIleValTyrTyrGlyGlyThrThrProThrGluArg 397
QY 731 TTGTTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTGCTCTCTA 790
Db 398 GlyIleSerAsnProTyrCysAla-----GlnLeuPheAlaSerArg 411
QY 791 GGTATGCTGTTGCTAGTATAGACACAGGGGATCCTGTCCAGGAGGCTTAAATTTGAA 850
Db 412 AspTyrValValTyrValIleGlnProSerGlyThrIleGlyPheGlyGlnGluPheSer 431
QY 851 GGCGCCTTTAAATATATAAATGGTCAATATAGAAATTTGACATCAGGTGGAAGGATCTCAA 910
Db 432 AlaArgHisValAsnAlaTrpGlyLysArgThrAlaAspAspIleIleGlyGlyThrLys 451
QY 911 TATCTAGCTTCTCGATATGATTTCATTGACTGATCGTGGGCATCCACGGCTGGTCC 970
```



! OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-11-264-096-1591

Alignment Scores:  
Pred. No.: 3.35e-11 Length: 349  
Score: 194.00 Matches: 37  
Percent Similarity: 69.4% Conservative: 13  
Best Local Similarity: 51.4% Mismatches: 22  
Query Match: 10.0% Indels: 0  
DB: 11 Gaps: 0

US-10-825-632-8 (1-1083) x US-11-264-096-1591 (1-349)

QY	263	GAAATATGATGATGCTGAAGAGGATCATAGTTGATGAGTCAGAGGCTGGTATAT	322
DB	129	GIUValLeuAlaArgHisGlySerLysIleTrpValAsnGluGluThrLysLeuValTyr	148
QY	323	TTTGAAGGCACCAAGACTCCCTTAGAGCATCACCTGACGTACGTACGTAAAT	382
DB	149	PheGlnGlyThrLysAspThrProLeuGluHisLeuTyrValValSerTyrGluAla	168
QY	383	CCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCATCTTCTGTCATCAGTCAG	442
DB	169	AlaGlyGluIleValArgLeuThrThrProGlyPheSerHis***CysSerMetSerGln	188
QY	443	CACTGTGACTTCTTATAGTATAGTAAACAG	478
DB	189	AsnPhe*****PheValSerHisIleThrAlaGln	200

Search completed: May 2, 2006, 03:10:58  
Job time : 38.1003 secs

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**



**THIS PAGE BLANK (USPTO)**